PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C12N 15/00, C12P 21/02
C12N 9/64, 15/00

A2

(11) International Publication Number: WO 88/ 05466

(43) International Publication Date: 28 July 1988 (28.07.88)

(21) International Application Number: PCT/US88/00109

(22) International Filing Date: 15 January 1988 (15.01.88)

(31) Priority Application Number: 003,611

(32) Priority Date: 15 January 1987 (15.01.87)

(33) Priority Country: US

(71) Applicant: CODON [US/US]; 213 East Grand Avenue, South San Francisco, CA 94070 (US).

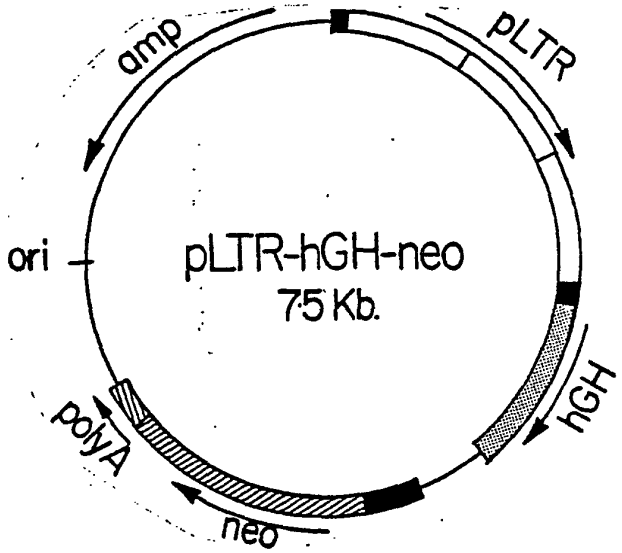
(72) Inventors: MORSER, Michael, John ; 3964 - 20th Street, San Francisco, CA 94114 (US). CASHION, Linda ; 219 Kelton Avenue, San Carlos, CA 94070 (US).

(74) Agent: WEBER, Ellen, L.; Townsend and Townsend, One Market Plaza, 2000 Steuart Tower, San Francisco, CA 94105 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: TANDEM GENE EUKARYOTIC EXPRESSION VECTORS



(57) Abstract

Compositions and methods for expressing gene products of interest in eukaryotic cells by transforming eukaryotic cells with recombinant DNA expression vectors which contain a DNA sequence coding for a gene product of interest together with a DNA sequence coding for a selectable phenotype. Both the gene of interest and the selectable gene are transcribed from a single eukaryotic promotor. Preferably, the recombinant DNA constructed is such that the DNA sequence coding for the gene product of interest is disposed between the eukaryotic promotor and the DNA sequence coding for the selectable phenotype. In an expression vector thus constructed, the gene product of interest is necessarily expressed by those cells which express the selectable phenotype. One clone of such transformed eukaryotic cells contains dicistronic mRNA which encodes for both human tissue plasminogen activator and the selectable phenotype. Another clone contains dicistronic mRNA which encodes for both erythropoetin and the selectable phenotype. Also disclosed are recombinant genes including an intron. The recombinant genes may be hybrids including both genomic DNA and cDNA.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

TANDEM GENE EUKARYOTIC EXPRESSION VECTORS

FIELD OF THE INVENTION

This invention relates generally to recombinant DNA expression vectors and more particularly to recombinant DNA eukaryotic expression vectors capable of expressing polypeptides in transfected eukaryotic cell lines. This invention also relates to a eukaryotic expression vector in which two genes are transcribed from a single promotor into a dicistronic mRNA.

BACKGROUND OF THE INVENTION

Numerous recombinant DNA techniques and expression vectors have been developed for use in polypeptide synthesis in prokaryotes, largely because the understanding in this area of bacterial genetics has proceeded so rapidly in previous decades. Because of the wealth of understanding and experience, the availability of a variety of advantageous prokaryotic expression vectors, and the ability to adapt bacterial systems to large-scale production, such systems have been preferred for the synthesis of various polypeptides of medical and pharmacological interests.

An example of such a production system is disclosed in U.S. Patent No. 4,342,832 which relates to the production of human growth hormone. The method disclosed therein, however, requires the inclusion of an ATG start signal codon on the 5′ end of the DNA structural gene. This produces a polypeptide with an amino-terminal methionine. When such a gene product is administered to a nonbacterial host, early clinical trials have indicated that an adverse immunological reaction may result because the methionine terminus may be recognized as foreign by the host cell. An additional problem with the production of eukaryotic gene products

in prokaryotic systems is that they may produce inactive products because abnormal eukaryotic protein folding may occur in a foreign prokaryotic environment.

It is now possible to introduce specific genes into cultured eukaryotic cells by various methods. It was originally found that by exposing eukaryotic cells to DNA which has been precipitated with Ca^{++}, a small fraction of the DNA molecules will stably integrate into the cell's chromosomal DNA. Graham, F.L. and van der Eb, A.J., Virology, 52:456-67 (1973). However, when such cells are transfected with genes which do not code for a selectable trait or phenotype, termed a "marker", the identification and expansion of such a transfected cell population is difficult.

Another procedure which has been used to transfect eukaryotic cells involves expression vectors based on the SV40 viral genome. However, the capacity of exogenous DNA which can be inserted in such a vector is severely limited due to the size of the viral capsid. (Old and Primrose, Principles of Gene Manipulation, 2d. Ed., University of California Press (1981)).

Another procedure, referred to as "co-transformation," is currently being used to transfect eukaryotic cells using physically unlinked sets of genes. This procedure selects for a subpopulation of cells competent in transfection, which cells are capable of integrating unlinked foreign genes into their genome at a higher frequency than the general population methods which involve the complementation of cell mutations by the co-transfer of selectable genotypes, however, are limited by the availability of mutant cell types which can serve as gene recipients. In addition when co-transfection is used to introduce both non-selectable and selectable genes, both plasmids may not integrate into the same region of the recipient cell's DNA and/or may not integrate into an active region of the chromosome. As a result, expression of the desired gene product in

the transfected cell may occur at relatively low levels, if at all.

Some transfection procedures have been developed for prokaryotes in which the vectors comprise a tandem arrangement of a plurality of genes. These procedures have not found much application in eukaryotic systems because normal eukaryotic mRNA's, unlike prokaryotic mRNA's, usually contain only a single cistron which is translated from a unique initiation site into a single protein. M. Kozak (Cell, 34:971-78 (1983)) has proposed that the eukaryotic ribosome scans the mRNA from the 5' end until it encounters the first ATG initiation codon at which protein synthesis initiation occurs. Translation continues until the first stop codon, at which translation terminates, is read and the ribosome disassociates from the mRNA, thus preventing the reading of other initiation codons. When more than one protein is encoded by a specific mRNA in a eukaryotic cell, protein processing of a precursor is generally responsible for generating multiple proteins.

Very recently, it has been found that some eukaryotic mRNA species contain more than one cistron, each with an independent initiation codon so that initiation and translation of proteins encoded by each cistron can occur. Multicistronic eukaryotic mRNA's, usually found in viral systems, may have a regulatory role in nature (Kozak, M. Cell 47: 481-483 (1986)), because the downstream genes are always expressed at lower levels (Peabody, D. and Berg, P. Mol. Cell Biol. 6: 2695-2703 (1986)). This property of differential expression of two tandem genes has not heretofore been utilized to design expression vectors that will overproduce the desired recombinant product relative to the selectable phenotype.

SUMMARY OF THE INVENTION

The present invention provides methods and compositions for the improved transfection of eukaryotic cells such that they express a desired gene product with greater efficiency and specificity than has heretofore been achieved. Such gene products usually take the form of the production of a foreign gene product which cannot be directly selected and may not be easily assayed, thus making clones producing such foreign products difficult to isolate. Such gene products include, but are in no way limited to, human growth hormone (hGH), human tissue plasminogen activator (t-PA) and erythropoetin (EPO).

The transfection of the present invention is accomplished by transfecting a eukaryotic cell with the novel eukaryotic expression vector of the present invention. The vector contains a structural DNA sequence coding for the desired gene product physically linked with a DNA sequence coding for a selectable phenotype which allows selection of the transfected cells. Because the two DNA sequences are physically linked to one another and depend on a common promotor, successfully transfected cells will contain the DNA sequences for both the desired gene product and the selectable phenotype. Thus, cells containing the DNA sequence coding for, and hence expressing, the desired gene product may be indirectly selected by selecting those cells which express the selectable phenotype.

The selectable phenotype is any one which allows selection of a subpopulation of transfected cells. Such selectable phenotypes are generally those which are not expressed in an untransformed cell but allow survival and growth of transfected cells in the presence of a particular agent. Such agents include, but are not limited to, aminoglycoside antibiotic G418, or methotrexate each of which is generally toxic to the untransformed cell population. The DNA of the transfected cells

contains DNA sequence coding for neomycin phosphotransferase (neo), dihydrofolate reductase (dhfr) and thymidine kinase (Tk), respectively. After transfection the cells are exposed to an environment containing the toxic agent in sufficient concentration such that those cells which do not express the selectable phenotype cannot survive. The surviving cells must, therefore, contain the DNA sequence coding for the selectable phenotype and the DNA sequence which codes for the desired polypeptide.

The transfecting vector of the present invention is generally comprised of a plasmid having a single eukaryotic promotor physically linked to a structural DNA sequence coding for the desired polypeptide and a DNA sequence coding for the selectable phenotype, followed by a signal for poly-adenylation. Preferably, the DNA sequence coding for the desired polypeptide is disposed between the eukaryotic promotor and the DNA sequence coding for the selectable phenotype. It is also preferred that the selectable phenotype is one that is not expressed in the untransformed cell. In one embodiment, the vector includes a plurality of promotor sequences of DNA. Further, the promotor includes eukaryotic enhancer sequences of DNA. Eukaryotic promotors and enhancers which have been shown to be useful include those derived from Harvey murine sarcoma virus and SV-40 virus. The vectors of the present invention are useful in a large variety of eukaryotic host cell lines including many laboratory cell lines.

The present invention also provides methods for synthesizing native polypeptides by transfecting eukaryotic cells with recombinant DNA of the present invention.

The present invention provides novel methods for producing a desired protein in a eukaryotic cell comprising the steps of:

(a) forming a eukaryotic dicistronic expression vector, which vector comprises:

(i) an intron-containing structural DNA sequence which codes for the desired protein, and

(ii) a second DNA sequence which codes for a selectable protein;

wherein the sequences are operably linked to one another and depend on one common promotor; and

(b) transfecting a eukaryotic cell with the vector and allowing this cell to grow under conditions favorable to the production of the selectable protein.

Preferably the promotor controlling the expression of the desired protein and selectable protein comprises a plurality of long term repeat regions of the Harvey murine sarcoma virus and the DNA sequence coding for the desired protein is operably disposed between the eukaryotic promotor and the DNA sequence coding for the selectable protein. Methods are further provided, in specific embodiments of this invention, whereby the recombinant DNA of the present invention results in the production of human tissue plasminogen activator or erythropoetin.

A method is further provided, according to the preceding paragraphs, wherein a portion of the DNA sequence coding for t-PA is derived from genomic DNA and contains at least one intron of between about 1-2kb. Preferably, the genomic portion of the DNA sequence coding for t-PA is a Bcl I - Bgl II fragment of the human t-PA gene containing a portion of the 5' untranslated region.

This invention further contemplates a method for producing t-PA in a eukaryotic cell compromising forming a eukaryotic expression vector that comprises a cDNA sequence and a genomic DNA sequence with at least one intron between about 1-2kb; and transfecting the cell with the vector. Preferably the genomic DNA sequence is a Bcl I - Bgl II fragment of the human t-PA

gene containing a portion of a 5' untranslated region and Intron A. Particularly preferred are the introns "A" to "F", using the designations of Ny et al., Proc. Nat'l Acad. Sci USA (1984) 81: 5355-5359 incorporated by reference herein, which describes the human tissue-type plasminogen activator gene and the exon-intron relationship. The sequence employed in connection with the coding sequence for t-PA will be characterized by having at least one intron between about 0.5 - 3.0kb more usually between about 1 and 2 kb and most preferably about 1.6 to 1.7 kb.

This invention further contemplates a method for producing EPO in a eukaryotic cell comprising forming a eukaryotic expression vector that comprises a structural DNA sequence encoding EPO which contains at least one intron of at least about 100 bp; and transfecting the cell with the vector. The sequence employed in connection with EPO will be characterized by having at least one intron between about 0.1- 3kb, more usually between about 0.1 and 1kb. Preferably the sequence will be characterized by having about 5-6 introns ranging in size from about 0.1 kb to 1kb.

This invention further contemplates eukaryotic cells which have been transfected by the methods described above, the recombinant DNA material produced by the above methods, and the vectors formed in the above methods.

The following terms are defined:

A "dicistronic vector" contains the DNA sequence of two genes that are operably linked such that they are capable of being transcribed from a single promotor into a mRNA promotor and such that a cell transfected by such a vector is capable of expressing the gene products from both of the genes.

An "intron-containing structural DNA sequence" is a DNA sequence which codes for a structural protein and contains at least one intron, but no more than about

13 introns, usually 2 to 8 introns. The introns which are employed may be naturally occurring introns associated with the sequence coding for the desired protein, or introns associated with the sequence coding for the desired protein or introns derived from a different mammalian gene. The precise placement of the introns may vary, but will typically be situated in the same location as found in the naturally occurring state.

A "desired protein" may be any of the variety of well-known recombinant proteins, particularly any of those having commercial value such as enzymes, hormones or those having antibiotic activity. Such proteins also include glycoproteins and lipoproteins.

A "selectable protein" is one that typically allows a cell to be selectively grown in a culture environment which otherwise would generally prevent growth of the cell.

"Operably linked" refers to expressible DNA sequences that are under the control of a promotor, e.g., such that a dicistronic mRNA is produced.

The manner in which these and other objects and advantages of the invention may be obtained will appear more fully from the detailed description of the invention which follows and from the accompanying drawings related thereto.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates the structure of the eukaryotic expression plasmid pLTR-hGH-neo. The following components have been incorporated into the plasmid: sequences from HaSV (filled in segments) including three LTR regions (open boxes); sequences coding for hGH (stipled box), sequences derived from Tn5 encoding the neo gene (hatched box), sequences from SV-40 containing the polyadenylation site (cross-hatched box), and bacterial plasmid vector pML (solid line).

Figure 2 illustrates the Northern blot analysis of cellular RNA from clone D4 containing the plasmid pLTR-hGH-neo, and described in detail below. Cellular RNA was prepared and analyzed as described. Lane 1 contains the cellular RNA hybridization with hGH probe; Lane 2 contains the cellular RNA hybridization with neo probe.

Figure 3 illustrates the characterization of cellular RNA from clone D4 and plasmid pLTR-hGH-neo as determined by S1 nuclease analysis.

a. DNA was labeled at the Bam HI, Bgl II or Sal I sites and then digested with Eco RI as indicated.

b. Cellular RNA was annealed to ^{32}P-labeled DNA fragments and then digested with S1 nuclease. The DNA fragments were fractionated on 1.5% agarose gel in 30 mM NaOH and identified by autoradiography. Lane 1 contains plasmid DNA labeled at the Bgl II site annealed to mRNA S1 digested; Lane 2 contains plasmid DNA labeled at Sal I site annealed to mRNA and S1 digested; Lane 3 contains plasmid DNA labeled at Bgl II site; and Lane 4 contains plasmid DNA labeled at Sal I site.

c. For DNA markers in this experiment, plasmid DNA was first labeled at the Bam HI site and then digested with Xba I, Sal I, or Kpn I. These three enzymes cleave within the LTR sequence (indicated by double line).

d. Plasmid DNA labeled at Bam HI site was hybridized with mRNA and then digested with S1 nuclease. The DNA fragments protected from S1 nuclease digestion were fractionated on a 5% acrylamide - 7M urea gel. The Bam HI labeled DNA fragment that was used as marker DNA was subcut with the following enzymes: Lane 1, Xba I; Lane, 2 Sac I; Lane 3, Kpn I. Lane 4 contains the annealed and S1 digested protected fragment.

Figure 4. Southern Blot analysis of cellular DNA from clone D4. Cellular DNA from untransfected C_{127}, (Lane A) and clone D4(Lane B) was digested with restriction enzymes as indicated and hybridized with a

^{32}P-labeled probe containing hGH sequences as described.

Figure 5 illustrates the analysis of supernatant proteins from clone D4. Supernatant proteins from clone D4 and C_{127} cells were labeled with ^{35}Smethionine and immuno-precipitated.

a. A comparison of immuno-precipitated and nonprecipitated proteins from the cell culture medium. Lane 1 contains ^{125}I-labeled standard hGH; Lane 2 contains clone D4 proteins precipitated with anti-hGH serum; and Lane 3 contains clone D4 proteins.

b. A comparison of the cytoplasmic proteins of clone D4 that have been immuno-precipitated with various antisera. Lane 1 contains ^{125}I-labeled standard hGH; Lane 2 contains ^{14}c- labeled molecular weight markers; lane 3 contains D4 proteins precipitated with anti-neo serum; Lane 4 contains D4 proteins precipitated with normal rabbit serum; Lane 5 contains D4 proteins precipitated with anti-hGH serum; and Lane 6 contains D4 proteins precipitated with normal mouse serum.

c. Partial chymotryptic digest of immuno-precipitated protein from the supernatant of clone D4. Lane 1 contains immuno-precipitated supernatant protein from clone D4; Lane 2 contains the same protein partially digested with chymotrypsin; Lane 3 contains authentic hGH labeled with ^{125}I; and Lane 4 contains commercially obtained hGH (Sigma Biochemical) partially digested with chymotrypsin. Figure 6 shows the construction of hGH expression plasmids. Two eukaryotic viral promotors were utilized from the hGH plasmids.

a. The plasmid containing the LTR promotor and the neo gene (pneo5) was obtained from M. Kreigler (Kreigler and Botchan, 1983). For pLTR-hGH-neo, sequences coding for hGH were excised from a cDNA clone (Martial et al, (1979)) with Hind III and inserted into a converted Bgl II site. The plasmid pLTR-hGHr-neo contains the hGH sequences in opposite orientation to the promotor. In order to delete two of the three LTR

segments, pLTR-hGH-neo was digested with Xba I which has one recognition site in each LTR, and the plasmid was religated (pLTR′-hGH-neo). The plasmid in which all three LTRS have been deleted (phGH-neo) was obtained as a spontaneous mutation during propagation in E. coli.

b. The SV-40 early promotor was contained in the plasmid pSV201-dhfr which was obtained from P. Berg (Subramani, Mulligan, and Berg, 1982). The hGH gene was inserted downstream from the SV-20 promotor at the Hind III site to generate pSV-hGH-dhfr. This plasmid was then modified to include the neo gene by an exchange of Bam HI fragments between pSV-hGH-dhfr and pLTR-hGH-neo.

Figure 7 illustrates the steps leading to the construction of transfection plasmids pPA003 and pPA005.

a. The DNA sequence of the amino terminal region of the t-PA gene extending from Intron A through the coding region and 5′ untranslated region to Intron A′. The dashed line shows the structure and sequence of the amino terminal region of the t-PA chromosomal gene as it is fused to t-PA cDNA to generate a hybrid genomic DNA-cDNA gene.

b. The Sequence of the synthetic "pre-pro" amino terminal region is shown. Synthetic DNA fragments were annealed to give 50 and 60 bp double stranded oligonucleotides with Bam HI or Bgl II and Nar I termini. The individual single stranded oligomers are indicated in boxes. The two double stranded segments were cloned separately in pUC9 that had been treated with BamH I and Nar I (a). These segments were removed by treatment of the recombinant pUC9 plasmids with Xho II and Nar I, the small t-PA DNA fragments were purified and ligated with pUC9 that had been treated with BamH I. This generated the entire pre-pro sequence. This sequence was excised from pUC9 by treatment with Xho II and cloned into the Bgl II site of the amino terminus of the t-PA cDNA to give plasmid pPA005 (b).

c. The Nar I fragment of pPA103 extending from intron A' to a Nar I site in t-PA cDNA was inserted into pPA104. This allowed removal of a 3.3 Kb Bcl I-Bgl II fragment containing the Bgl II site. This fragment was ligated into pneo5 resulting in pPA003. the genomic hybrid t-PA expression plasmid.

d. This Figure illustrates plasmid pPA005. This t-PA expression plasmid was made from synthetic DNA encoding the pre-pro region ligated to the cDNA encoding mature t-PA.

Figure 8 shows the results of Northern blot analysis on poly-adenylated mRNA taken from cells transfected with the pLTR-tPA-neo plasmids of Figure 7.

Figure 9 illustrates the Northern blot analysis of cellular RNA from clones derived from CHL-1 cells transfected with pPA003. The probe used in this analysis was the 1.6kb Bgl II fragment containing t-PA cDNA.

Figure 10 illustrates the complimentarity of cellular RNA from clone M5A6 with the LTR promotor. The plasmid pneo5 was digested with Eco R1 and Bgl II, the fragment was isolated, and then labeled with polynucleotide kinase. The labeled fragment was annealed with total RNA from clone M5A6, and then digested with S1 nuclease. The DNA fragments protected from S1 digestion were fractionated on a 5% acrylamide - 7M urea gel. Viral DNA from X174 RF digested with Hae III and labeled with polynuclectide kinase was used for size markers.

Figure 11 illustrates that the Northern blot analysis of clones generated in CHL-1 cells. Clones where isolated following transfection with the following plasmids indicated in parentheses: pPA007 (73-1), pPA002 (73-2), pPA003 (73-4, 73-5, and M5A6) and pPA004 (73-6, 73-7).

Figure 12 illustrates plasmid pPA201. The blackened area represents the SV-40 promotor; the dashed segment represents the dhfr gene; the striped segment

represent the introns of the t-PA genomic fragment, the white areas within this segment indicate the portions of the genomic fragment that contain t-PA exons; the dotted region represents the cloned t-PA cDNA fragment. The arrows indicate the direction of mRNA transcription.

Figure 13 illustrates plasmid pPA401. The blackened area represents the triple LTR promotor; the dashed segment represents the tk gene; the white region represents the synthetic cDNA 5' region of t-PA constructed from oligoneucleotides; the dotted region represents the cloned t-PA cDNA fragment. The arrows indicate the direction of mRNA transcription.

Figure 14 illustrates the plasmid pEP1 showing the relationship of the LTR promotor to the coding sequences of EPO and neo. The sequences encoding EPO are contained in a 4kb Apa I genomic fragment. The introns within this region are indicated by the numbered dotted segments. The exons are represented by the blackened areas. The amino acids encoded by each exon are indicated. The arrows indicate the direction of mRNA transcription.

Figure 15 shows plasmid pEP2 indicating the relationship of the LTR promoter (the blackened area) to the coding sequences of EPO gene (the striped region) and the neomycin resistance gene (dashed region). The sequences encoding EPO are derived from synthetic DNA oligomers based on the sequence of EPO cDNA. The arrows indicate the direction of mRNA transcription.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides methods and compositions for the transfection of eukaryotic cells with a novel expression vector containing a DNA sequence which codes for a desired polypeptide and a selectable phenotype not expressed in the untransformed cell. Thus, utilizing the present invention the transfection of eukaryotic cells may be monitored, and thereafter,

successfully transfected cells may be assayed for their ability to express the selectable phenotype. In one embodiment, the desired polypeptide is not expressed in the untransformed cell. In another embodiment, the desire polypeptide is one that is already expressed in the untransformed cell but whose expression is enhanced after transfection. Embodiments of the present invention include those in which the gene coding for the desired polypeptide codes for human growth hormone, tissue plasminogen activator or erythropoetin. Although the specific embodiments described herein involve cultured mammalian cells, it may be appreciated that the methods and compositions of the present invention are generally applicable to all eukaryotic cells including plants and micro-organisms, such as yeast.

The present invention utilizes a vector for transfection of eukaryotic cells that comprises a eukaryotic promotor directly coupled to a DNA sequence coding for a desired polypeptide and a DNA sequence coding for a selectable phenotype. In a preferred embodiment, the DNA sequence coding for the desired polypeptide is disposed between the eukaryotic promotor and the DNA sequence coding for the selectable phenotype.

The eukaryotic promotors which have been used in the transfection vectors of the present invention include the early promotor ("SV") of the SV-40 virus and the long terminal repeat ("LTR") region of the Harvey murine sarcoma virus ("Ha-MuSV"). Plasmids containing these promotors are hereinafter designated "pLTR-" and "pSV'," respectively.

The eukaryotic promotors used in the vector of the present invention also include enhancer sequences which increase rates of RNA transcription presumably by facilitating the binding of RNA polymerase.

Since proper expression of a bacterial gene in a eukaryotic cell requires polyadenylation on the 3'-end of the mR NA, the vectors of the present invention

also contain appropriate sequences responsible for adding a poly-A tail following the gene coding for the selectable phenotype.

According to the present invention, plasmids have been prepared which can be propagated in prokaryotic cells by virtue of a selectable marker, e.g. an antibiotic resistance gene. The plasmid further includes a gene coding for a selectable phenotype which allows eukaryotic cells transfected with the plasmids of the present invention to be selectively grown in the presence of an agent which otherwise generally prevents growth of the eukaryotic cells. The gene coding for the selectable phenotype in the eukaryotic cell to be transfected may be the marker that allows selection of the plasmid itself or another plasmid. P.J. Southern and P. Berg (J. Mol. App. Gen., 1:327-41 (1982)) have described the selection of antibiotic resistant clones of transfected eukaryotic cells.

The DNA sequences coding for selectable phenotypes which have been successfully employed in the vectors of the present invention include those coding for neomycin phosphotransferase ("neo"), dihydrofolate reductase ("dhfr") and thymidine kinase ("TK"). The neo selection marker may be derived from bacterial transposon TN5 which confers resistance to kanamycin in bacteria and to the aminoglycoside antibiotic G418 ("G418") eukaryotes. Southern, P.J. and Berg, P. supra. The dhfr gene confers resistance to methotrexate ("MTX") by the expression of dihydrofolate reductase. The TK gene is derived from the TK gene of herpes simplex virus. Expression of this gene results in the production of thymidine kinase. When the gene is present in a cell that is otherwise incapable of producing thymidine kinase, that deficient cell, once transfected, may survive in the presence of thymidine due to the production of exogenous thymidine kinase.

Phenotypes of interest can take the form of the expression of any gene product. Although these may include gene products which the cell is capable of producing in its untransformed state, they will generally include those gene products which are foreign to the cell. Because of the tandem arrangement of genes physically linked to a single eukaryotic promotor in the plasmid of the present invention, cells which grow in the presence of the agent must also express the gene coding for the desired polypeptide. In one embodiment of the present invention, a gene coding for hGH is inserted in the plasmid at a site located between the promotor sequence and the selectable marker gene. In another embodiment, a gene coding for t-PA is inserted between the promotor sequence and the selectable marker gene. In yet another embodiment, a gene coding for EPO is inserted in the plasmid at a site located between the promotor sequence and the selectable marker gene.

The particulars of the construction and use of the eukaryotic expression vectors of the present invention are more fully disclosed by way of example as shown below:

Experimental

The following examples illustrate the usefulness of the methods and compositions of the present invention. The relative efficiencies of the methods and compositions of the present invention may be compared to co-transfection methods and vectors. The following examples also apply the methods and compositions of the present invention to various desired polypeptides, selectable phenotypes, promotors and eukaryotic host cells. In the following examples, the desired polypeptides include human growth hormone, human tissue plasminogen activator and erythropoetin, although as one of ordinary skill in the art will appreciate any desired

polypeptides having a known or identifiable gene may be expressed.

All of the DNA transfection examples which follow were carried out using the $CaCl_2$, precipitation technique of F.L. Graham and A.J. van der Eb (Virology, 52:456-67 (1973)). The transfection protocol followed in these examples are as follows: Transfection Protocol

Day 1 Plate out 2 x 10^5 cells/well in 6 well plates in DME + 10% FBS (fetal bovine serum).

Day 2 Carry out transfection:

1. DNA is diluted to 50 ug/ml (micrograms per milliliter) in 100 ul (microliters) H_2O. Five micrograms of DNA will be needed to transfect 2 x 10^5 cells

2. 2M $CaCl_2$, (0.125 ml) is diluted to 0.25 M in hepes-buffered saline 0.25 (HBS) (0.875 ml).

3. DNA and $CaCl_2$, (0.25 M) are mixed 1:1. A precipitate is allowed to form at room temperature for 30 min. - 1 hr.

4. The medium is removed from the cells in the well plate, and the precipitated DNA is added. The DNA is absorbed into the cells for 20 min, and then diluted with 10 volumes of fresh medium.

5. The cells are incubated further for four hours in the diluted DNA solution.

6. After four hours, the medium containing the DNA is removed, and the cells are treated with 25% glycerol in HBS for one minute. The cells are washed once with medium, and then 3 ml of medium containing 10% FBS is added to the cells.

Day 4 Transfer cells to selective medium.

1. The cells in each well of the 6 well plates are trypsinized and counted.

2. The cells are diluted to yield 10^3, $3 \cdot x$ 10^3 and 10^4 cells/ml. The selective agent is also added to the diluted cells. For each cell line, the amount

of the selective agent required to completely kill the transfected cells must be determined.

After the cells have been plated into the 96 well plates, the medium is changed twice weekly until the clones are visible. For most selection procedures, clones are visible after three weeks in the selective medium and cells can be assayed four weeks after transfection.

Transfection frequencies for the various plasmids and host cells of the following examples are determined by counting the numbers of clones in a 96 well plate and dividing that number by the total number of cells in the plate.

Human Growth Hormone

In the following examples, the phenotype of interest is the expression of human growth hormone ("hGH").

Human growth hormone has been used in the past to treat children who are deficient in the production of hGH and suffer from dwarfism. Previously, the only source of purified hGH was from human cadavers, and the quantity produced was not sufficient the demand for the product. More recently it has been found that some of the early preparations of hGH were contaminated by a fatal virus that undergoes a long latency period, and therefore, the use of these preparations of hGH has been discontinued. The only source of hGH now available for treatment of hGH deficient dwarfism is the hGH made in E. coli that contains the N-terminal methionine. At this time, it is not known if the use of the modified hGH will lead to eventual immunological problems for these children later in life.

For each transfection example for hGH described below, 5ug of transfection plasmid DNA was added to a monolayer of approximately 10^5 host cells in 35mm dishes. Of the following examples, the selectable phenotype in Examples 1-8 is one that confers resistance to G418. None of the eukaryotic host cells were resistant to

G418 in their untransformed state. Clones which demonstrated resistance to G418 at a concentration of 400 ug/ml were isolated and identified 2 to 3 weeks following transfection. All G418 resistant clones were screened for hGH by radioimmunoassay techniques (Pharmacia Fine Chemical, Piscataway, New Jersey).

The results of the experimental transfections relative to each of the desired polypeptides are summarized in Table I which follows Example 9 below.

Example 1

Transfection plasmid pLTR-hGH-neo is shown in Figure 1. Sequences coding for hGH were inserted into the base plasmid pneo-5 as described below. The construction of the plasmid pneo-5 has been previously described by Lusky, M. and Botchan, M. (Cell 36:391-401 (1984)). The eukaryotic promotor from Ha-MuSV was inserted in bacterial plasmid pML to form pML-LTR. The promotor for viral proteins in this plasmid is located at both ends of the DNA copy of viral RNA and is contained in the long terminal repeat ("LTR") regions. To construct the pML-LTR plasmid, the LTR region was excised from the plasmid pRetro-T III with enzymes Cla I and Bam HI. The LTR region of p-Retro-T III contains three copies of the LTR promotor. This excised LTR region was inserted into the bacterial plasmid pML (Lusky, M. and M. Botchan, Nature, 293:253-58 (1981)) to form pML-LTR. The TN5 element, containing the neo gene, was obtained from a SV40 hybrid plasmid. This TN5 element was removed from the SV40 hybrid by digestion with Bgl II and Bam HI. The TN5 fragment thus obtained was then inserted into the Bam HI site of pML-LTR and recombinant plasmids were then screened for the proper orientation of the TN5 element with respect to the LTR promotor. A poly A addition sequence of the SV40 T antigen was excised from SV40 DNA by digestion with Bcl I and Bam HI and inserted in the Bam HI site of the above plasmid described above forming plasmid pneo-5.

Plasmid phGH800/pBR322 which contains an hGH gene bounded by Hind III restriction endonuclease sites, served as the source of the hGH gene for the transfection plasmid of this example. This plasmid is described in Martial, J.A., Hallewell, R.A., Baxter, J.D., and Goodman, H.M. Science 205:602-606 (1979) and was obtained from this group. The hGH gene of this plasmid was excised by digestion with Hind III. The Bgl II site on plasmid pneo-5 was converted into a Hind III site, and the hGH fragment was inserted into the converted Hind III site of plasmid pneo-5 to form transfection plasmid pLTR-hGH-neo. The recombinant clones were screened by restriction enzyme digestion to determine the proper orientation of the hGH genes following the LTR promotor.

The pLTR-hGH-neo plasmids that showed proper orientation were used to transfect 10^5 C_{127} cells, a mouse embryonic breast cell line. Radio immunoassays demonstrated that 100 percent of the clones which grew in the presence of G418 also expressed significant levels of hGH, ranging from 0.01 to 0.4 pg(picograms)/cell/day. (See Table I). Thus, it appears from these results that two genes can be transcribed from a single promotor in these clones. The pLTR-hGH-neo plasmid and one clone from the above example (Clone D4) were deposited in the American Type Culture Collection (ATCC), Rockville, Maryland for the required period of 30 years, as follows:

JM83/pLTR-hGH-neo was deposited on February 16, 1984 and assigned ATCC accession number 39614; and

C_{127}/pLTR-hGH-neo was deposited on February 16, 1984 and assigned ATCC accession number CRL-8503.

A Northern analysis was performed on the mRNA from the clone D4, generated by pLTR-hGH-neo into C_{127} cells, to determine how this mRNA was transcribed. Total RNA was isolated from clone D4. The mRNA was purified by batch elution from oligo-dT cellulose Poly-A containing mRNA was fractionated by electrophoresis

through a 1% agarose-formaldehyde gel to denature the mRNA (Maniatis, T., Molecular Cloning, Cold Spring Harbor, New York (1982)) and then transferred to nitro-cellulose. Hybridization with 32P-labelled probes was carried out using the method of Wahl, G.M. et al (Pro. Natl. Acad. Sci USA, 76:3683 (1979)). Hybridization was carried out with a hGH-specific probe whereupon a single band of complementary mRNA was detected with an apparent size of 3.5 kb (Figure 2). The DNA probe previously annealed to mRNA was removed by immersing the nitro-cellulose filter in boiling water and rehybridizing with other probes. Upon hybridization with a TN5-specific probe, an mRNA of identical size was found (Figure 2). Using an agarose-urea gel system, the apparent size of the mRNA species from the transfected cells was verified. The DNA sequences coding for both hGH and TN5 are contained within a single 3.0 kb fragment of DNA approximately 300 bp downstream from the LTR. This data suggests that the mRNA with a size of 3.5 kb is the only species of hGH or TN5 mRNA within the cells transfected with pLTR-hGH-neo because mRNA processing did not give rise to two different mRNA species, one coding for hGH and the other coding for neomycin phosphotransferase.

An S1 nuclease analysis was then carried out to confirm that both the hGH gene and the neomycin phosphotransferase gene are encoded in the same mRNA and to discover where within the LTR promotor of pLTR-hGH-neo the initiation site of the mRNA transcript was located. The S1 protection assays followed the procedure of Maniatis, supra (1982) and is described generally below.

DNA from pLTR-hGH-neo was digested with either Bam HI, Bgl II, or Sal I, which cleave either in the promotor, hGH or TN5 sequences (see Figure 3a). The digested DNA was end-labelled with ^{32}P (New England Nuclear) and polynucleotide kinase (P-L Biochemicals). The DNA was then digested with Eco R1 and the appropriate fragments isolated from a 5% polyacrylamide gel.

Approximately 100 ug of DNA was annealed to 2 ug mRNA at 50 degrees centigrade for 3 hours and the non-complementary nucleic acid was digested with S1 nuclease (Miles Laboratory). The protected DNA fragments were electrophoresed through a 1.5% agarose-alkaline gel or a 5% acrylamide gel containing 8 M urea. (Maniatis, supra (1982)). The results are shown in Figures 3b and 3c.

A single band at 0.3 kb was detected when DNA fragments labeled at the Bam HI site were hybridized to the cellular mRNA. When Bgl II digested DNA was hybridized to the cellular mRNA, a band was detected at 0.8 kb. Finally, when Sal I digested DNA is hybridized to the mRNA, a band appeared at 3.0 kb (Figure 3b). No splicing of the mRNA was observed by this technique. It appears from the results of this S1 analysis, that transcription of the mRNA begins at a unique site within the LTR promotor, 300 bp upstream from the hGH gene and terminates within the SV40 fragment containing the poly-A addition site which follows the TN5 gene (Figure 3a). Therefore, the results of the S1 analysis further support the conclusion that there is only a single species of mRNA that encodes both hGH and the neomycin phosphotransferase genes in the cells of clone D4.

The results of the mRNA analysis suggest that the gene sequences for hGH and neo are integrated together in one expression unit. To test this hypothesis, DNA was isolated from clone D4 and analyzed by Southern blot analysis (Southern, E.J., Mol. Biol. 98:503-518 (1979)). The DNA was digested with Bam HI, Cla I, and Pst I which cleave the plasmid pLTR-hGH-neo to yield hGH containing fragments of 3.0, 6.0 and 1.6 kb, respectively (Figure 4). Following hybridization of the digested clone D4 DNA with an hGH probe, bands of the expected sizes were observed. Other bands are also observed which are probably due to fragments of plasmid DNA that have integrated.

The protein product of clone D4 in the supernatant of the culture identified by radioimmunoassay was characterized to determine its moleclar weight. To this end, cells of clone D4 were radioactively labelled with ^{35}S-methionine. The supernatant proteins secreted into the incubation medium by these cells were immunoprecipitated with mouse anti-hGH serum. The resultant hGH material was fractionated on 15% SDS-polyacrylamide gels according to the method of U.K. Laemmli, (Nature, 227:680-85 (1970)). As seen in Figure 5a, the immunoreactive protein product of clone D4 co-migrates with naturally synthesized hGH at a molecular weight of 22,000 (Figure 5a, lanes 1 and 3). A prominent band of 22,000 molecular weight can also be seen in the total supernatant protein (Figures 5b, lanes 1 and 2). This hGH like protein is approximately 3% of the supernatant. No similarly sized protein was found in control cells which have not been transfected with pLTR-hGH-neo (Figure 5b, lanes 5 and 6).

A partial chymotryptic digest of the 22,000 molecular weight hGH immuno-reactive protein product of Clone D4 was performed according to the method described by Cleveland, D. W. et al. (J. Biol. Chem. 252:1102-1106 (1977)) to determine whether this protein product was processed in a manner conforming with naturally synthesized hGH. Due to its native conformation naturally synthesized hGH is cleaved to smaller peptides of 15,000 and 7,000 when subject to partial proteolysis. The 15,000 peptide is derived from the N-terminal region of the protein. (Aston, R., EMBO. J., 2:493-97 (1983)). Cells transfected with pLTR-hGH-neo were labeled with 35S-methionine and the supernatant proteins were precipitated as described in the preceding 35S- methionine assay. Ten mg of the immunoreactive protein was partially digested with 10ug of chymotrypsin for 30 minutes at 37 degrees centigrade according to the method of Cleveland et al., supra (1977). The products were

compared with ^{125}I-labelled partially digested standard hGH protein by fractionating both digestions on a 15% polyacrylamide SDS gel (Laemmli, supra (1970)). The results demonstrate that both digests contain the 15,000 MW partial peptide derived from the N-terminus region of the protein (Figure 5c). Since the partial chymotryptic maps of both digests were identical, it appears that the hGH secreted from transfected cells has been translated and properly processed at the N-terminus. Thus, clones which are resistant to G418 and which also express a 22,000 molecular weight protein immuno logically related to hGH can be isolated from cells transfected with pLTR- hGH-neo.

Example 2

The construction of the pLTR-hGH-neo transfection plasmid of Example 1 resulted in some plasmids in which the hGH DNA sequence was inserted in an opposite orientation to the promotor to form pLTR-hGHr-neo. One such plasmid was used to transfect C_{127} cells as described in Example 1. All other procedures of Example I were followed. Although the transfection efficiency remained unchanged from that of Example 1, none of the clones expressed hgH (Table I).

Example 3

The expression vector of Example 1 contained three adjacent copies of the LTR promotor. Two of the three LTR promotors were deleted by digestion with Xba I and the plasmid relegated to pLTR′-hGH-neo plasmid. Plasmid pLTR′-hGH-neo was used to transfect $^{C}127$ cells as in Example 1: All other procedures of Example 1 were followed. However, the transfection efficiency of this example was reduced ten-fold as compared to that of Example 1 and the frequency of clones expressing hGH was reduced to 25% as compared to 100% for pLTR- hGH-neo of Example 1 (Table 2).

Example 4

All of the LTR promotors of the plasmid of Example 1 were deleted to form phGH-neo. This plasmid was used to transfect C_{127} cells as described in Example 1. Even further decreases in transfection efficiency and frequency of hGH expressing clones occurred when the entire LTR promotor region was deleted as compared to the results of Example 3.

The results of Examples 2, 3 and 4 indicate that the expression of hGH is mediated by the LTR promotor. Examples 3 and 4 further indicate that expression of both hGH and neo is greatly enhanced when more than one copy of the LTR sequence is available in the eukaryotic promotor of the transfection plasmid of the present invention.

Example 5

The pLTR-hGH-neo transfection plasmid of Example 1 was used to transfect COS-7 cells, a green monkey kidney cell line. Clones were selected by their ability to grow in the presence of G418. Six G418 resistant clones were produced of which 3 were screened for hGH expression levels. The clones producing hGH had hGH expression levels ranging from 0.6 to 1.3 pg/cell/day.

Example 6

The pLTR-hGH-neo transfection plasmid of Example 1 was used to transfect CHL-1 cells, a human melanoma cell line. Clones were selected by their ability to grow in the presence of G418. The G418 resistant clones numbered 8, of which 5 were screened for hGH expression levels. The clones producing hGH had hGH expression levels ranging from 0.02 to 0.64 pg/cell/day.

Example 7

As a comparative study, a co-tranfection system employing two plasmids was used to transfect C_{127} cells. Plasmids containing a eukaryotic promotor and the hGH gene were used together with physically unlinked

plasmids containing a eukaryotic promotor and a selection marker gene in the cotransfection procedure.

In the construction of plasmids containing an SV-40 promotor, the promotor and dhfr coding sequences from pSV2dhfr (Subramani, Mulligan, and Berg, Mol. Cell Bio., 1:854-64 (1981)) were transferred to pBR328 to eliminate eukaryotic poison sequences (Lusky and Botchan, 1981). This new plasmid (pSV-dhfr) was linearized with Hind III, and the hGH sequences were inserted between the promotor and dhfr gene (Figure 6). The hGH of pBR322 gene was inserted downstream from the SV promotor of the pSV-dhfr plasmid and in front of the sequences coding for dihydrofolate reductase at the Hind III site to generate pSV-hGH-dhfr.

Although transfection efficiency of this example was similar to the efficiency obtained by the method of Example 1, radioimmunoassay revealed that only 13.6% of the clones obtained by co-transfection expressing the selectable phenotype also expressed hGH (Table 1).

Moreover, hGH expression levels of clones obtained by cotransfection were all significantly lower than clones generated from tandem expression vectors of the present invention. The G418 resistant clones obtained by this co-transfection method expressed hGH in the range of .002 to .03 pg/cell/day whereas the G418 resistant clones of Example 1 expressed hGH in the range from .01 to 1.0 pg/cell/day.

Example 8

An alternative embodiment of the present invention utilizes the early promotor from SV40 virus derived from the pSV201-dhfr plasmid as described in Subramani et al., Mol Cell Biol., supra (1981). A schematic representation of the construction of this plasmid is shown in Figure 6. The SV40 early promotor is known to function effectively in a wide variety of eukaryotic cell types with the prokaryotic TN5 gene. To insert

the hGH into the SV-40 promotor plasmid, pSV-hGH-dhfr and pLTR-hGH-neo were digested with Bam HI, and the Bam HI fragment containing sequences for both hGH and neo was inserted downstream from the promotor (pSV-hGH-neo).

The resultant pSV-hGH-neo plasmid was transfected into $^{C}127$ cells and clones were selected by their ability to grow in the presence of G418. When the selected G418 resistant clones were tested, it was found that 100 percent of the clones produced hGH in the range of 0.002 to 1.00 pg/cell/day. (See Table 1.)

Example 9

As an alternative comparative co-transfection system, hGH producing clones were identified in a human cell line. On clone of this human cell line CHL-1 was generated by co-transection using pSV-hGH-dhfr described in Example 8 and pneo-5 described in Example 1. This clone was then grown in the presence of 2 μM methotrexate ("MTX"), a folic acid antagonist. The presence of MTX stimulated an overproduction of dhfr and consequently, since these genes were linked, the production of hGH was increased ten-fold as compared to the co-transfection of Example 6. Thus, it is possible to amplify the expression of the gene of interest that is physically linked to and shares the same promotor as the selective marker.

Human Tissue Plasminogen Activator

Another embodiment of the present invention involves the production of human tissue plasminogen activator ("t-PA"), a pharmaceutically important thrombolytic agent.

For each transfection example for t-PA described below, 5ug of transfection plasmid DNA was added to a monolayer 10^{6} host cells in 100 mm dishes. In Examples 10-13, the selective phenotype confers resistance to G418 at a concentration 1.0 mg/ml. Clones which demonstrated a resistance to 1.0 mg/ml G418 were isolated and identified 3 to 4 weeks following

transfection. In Example 14, the selectable phenotype is the expression of dhfr which confers upon the transfected cell the ability to grow in the absence of hypoxanthine, glycine and thymidine and in the presence of methotrexate. In Example 15, the selectable phenotype is the expression of thymidine kinase and the transfected cell the ability to grow in the presence of methotrexate.

When the isolated clones contained at least 10^3 cells, the clones were screened for t-PA expression. Screening was determined via fibrin plate analysis for a 24 hour period. For this analysis, 5 ul samples of culture supernatant were placed in a circular well excised from an agarose plate containing sheep fibrogen, human thrombin and human plasminogen. A clearing around the circle of culture supernatant was caused by conversion of plasminogen to plasmin which had digested the fibrin, all initiated by the presence of t-PA. The size of the clearing, thus, correlates with the amount of t-PA in the sample. To determine the t-PA expression levels, the amount of t-PA in the culture supernatant was divided by the number of cells in the culture. The results of these examples are summarized in Table V below.

Example 10

Sequences coding for the protein tissue plasminogen activator (t-PA) were inserted into the pneo5 plasmid in a manner analogous to that used to insert hGH into pneo5 of Example 1. The DNA sequences coding for t-PA were obtained from a human fetal liver DNA library and also from a complementary DNA transcribed from CHL-1 cell mRNA.

The isolation and assembly of the DNA sequence encoding the t-PA gene is depicted in Figure 7. CDNA was produced from mRNA derived from the melanoma cell line CHL-1. This cDNA was digested with Bgl II and the fragments were cloned into a Charon 4A vector at the Bam HI site (Maniatis et. al. Molecular Cloning). A

TABLE I

Example	Plasmid	Host Cell	Agent	Transfection Efficiency (x10-3)	hGH+clones/ Total Analyzed	Expression Levels (pg/cell/day)
1	pLTR-hGH-neo	C_{127}	G418	1.4	10/10	0.01-0.4
2	pLTR-hGHr-neo	C_{127}	G418	1.4	0/10	------
3	pLTR'-hGH-neo	C_{127}	G418	0.15	2/8	0.02
4	hGH-neo	C_{127}	G418	0.7	2/18	0.0002
5	pLTR-hGH-neo	COS-7	G418	0.03	3/3	0.6-1.3
6	pLTR-hGH-neo	CHL-1	G418	0.03	5/5	0.2-0.6
7	pSV-hGH-dhfr+ pneo5	C_{127}	G418	2.6	3/22	.01-03
8	pSV-hGH-neo	C_{127}	G418	2.4	22/22	0.02-1.0
9	pSV-hGH-dhfr+ pneo5	CHL-1	G418	0.06	8/22	0-0.3

subclone containing the Bgl II fragment which extends from bp187 to 2161 of the coding sequence of t-PA was used in the construction of t-PA expression vectors. This fragment encodes the mature form of the t-PA protein but does not contain sequences for the pre- and pro- peptides of t-PA required for secretion of the active protein.

The Bgl II fragment from bp 187 to 2161 was removed from the Charon 4A clone and subcloned into pUC19. Sequences extending only from the Bgl II site at bp 187 to the Xho II site at 1805 were further subcloned by Xho II digestion of this pUC19 derivative and were inserted into the Bam HI site of pUC9. Xho II digestion of the pUC9 derivative allowed retrieval of the cDNA fragment. The Xho II fragment was subcloned into the cloning vector p31 at its Bam HI and Bgl II sites to produce pPA106. The Xho II fragment was also subcloned into p31 at its Bgl II site only to produce pPA104.

In order to express and secrete the mature and active form of t-PA in transfected cells, it was necessary to provide sequences coding for the secretory process. The sequences for the pre- and pro- peptides were obtained by two different methods. In the first case a genomic fragment was isolated from a human genomic DNA library (Figure 7a). These genomic sequences were ligated in front of the cDNA to produce a hybrid t-PA gene. In the second case the front of the t-PA gene was synthetically produced from small oligonucleotides (Figure 7b).

The specific genomic sequences encoding the pre- and pro-peptides were isolated from a πVX library. This library was constructed using the 414bp Pst I fragment from the t-PA cDNA as the target sequence on πVX (Manniatis et al. supra). One clone was found which contained a 4kb Bgl II fragment encompassing the 5' end of the t-PA structural gene including 105 nucleotides

of the pre- and pro- peptides, a portion of the 5′ untranslated sequences and a large intron (Figure 7c). This 4kb Bgl II fragment was subcloned into pPA106 at its unique Bgl II site to generate the subclone pPA103. The cDNA/genomic junction was removed by digestion at the Nar I site in the cDNA and a Nar I site in the genomic portion. This Nar I fragment was then moved into the other cDNA subclone, pPA104, at its equivalent Nar I site (Figure 7c). The new subclone, pPA115 now contained the entire genomic and cDNA sequence. The hybrid t-PA gene was removed by digesting with Bcl I and partial digestion with Bgl II. The Bcl I - Bgl II cassette was then inserted into the unique Bgl II site of pneo5 thus qenerating pPA003. Because the t-PA coding region contains both genomic (including exons and introns) and cDNA sequences, this construction is referred to as a hybrid t-PA gene. The Bcl I-Bgl II fragment is illustrated in Figure 7a. This fragment encompasses several bases (about 32) from intron A′ (as labelled on Fig. 7a and 7c), a 5′ untranslated region, the pre- and pro-peptide region and a large intron A of about 1.7kb. Intron A′ described herein corresponds with the intron "A" designation of Ny et al., supra. Intron A described herein corresponds with the intron "B" designation of Ny et al., supra.

Sequences encoding the pre- and pro- peptides of t-PA but lacking the intron found in genomic DNA were also assembled from synthetic DNA (Figure 7b). Synthetic DNA fragments were annealed to give 50 and 60 bp double stranded oligonucleotides with Bam HI or Bgl II and Nar I sticky ends. These two double stranded segments were cloned separately in pUC9 that had been linearized with Bam HI and Nar I. These cloned segments were subsequently removed by treatment of the recombinant plasmids with Xho II and Nar I, purified and ligated to pUC9 that had been treated with Bam HI. This generated the 110 bp pre-pro sequence. This 110 bp sequence was

excised from pUC9 by treatment with Xho II and then cloned into the Bgl II site of pPA024 to form plasmid pPA005 which contains a complete cDNA gene (Figure 7d).

Both of these t-PA expression plasmids, pPA005 and pPA003, were used to transfect CHL-1 cells and clones derived from both plasmids have been analyzed. Since this cell line produces endogenous t-PA, it was necessary to screen all G418 resistant clones for increased levels of t-PA production to determine if exogenous t-PA was also produced.

Additional plasmids were also constructed in which the triple LTR promotor at pPA005 was deleted and other eukaryotic viral promotors were inserted. The plasmid pPA014 was obtained from pPA005 by deletion of the triple LTR promotor. In one case, the promotor from the herpes simplex virus thymidine kinase gene was inserted into pPA014 to generate pPA002. Two other promotors from Adenovirus 5 were substituted for the triple LTR promotors (pPA007 and pPA008).

TABLE II

Transfection Frequencies of t-PA Expression Plasmids

Plasmid	Exp.#	Clones/10 cells
pPA002	1	11
	3	11
	4	9
pPA003	1	55
	2	22
	3	31
	4	24
pPA005	2	45
	3	26
	4	38
pPA007	1	15
	2	21
pPA008	1	8
	2	13
pPA014	3	15

The above plasmids were used to transfect CHL-1 cells in culture using the protocol described above for hGH plasmids. A summary of the transfection results is given in Table II for each plasmid, in which the number of clones per 10^5 cells was determined. In each case, the transfection frequency for the t-PA plasmids was lower than that seen for the hGH plasmids. Furthermore, the plasmids containing the triple LTR promotor (pPA003 and pPA005) gave slightly more clones. Most of the clones obtained from these experiments described in Table II were assayed for t-PA activity. The highest t-PA producing clones were maintained in medium lacking the selective agent and then assayed further to compare expression levels. The results of these assays are given in Table III. The clones expressing the highest levels of t-PA in the absence of the selective agent were obtained from plasmids containing the triple LTR promotor (pPA003 and pPA005). Clones obtained from other plasmids were not expressing t-PA at levels greater than clones obtained from pPA014 which lacks a promotor. Moreover, the plasmid containing the hybrid t-PA gene generated clones producing the highest levels of t-PA.

A further comparison between the hybrid t-PA gene assembled in pPA003 and the t-PA gene containing the 110 bp synthetic fragment pPA005, is shown in Table IV. In Table IV, the expression levels of the highest clones derived from pPA003 and pPA005 from several experiments are compared. From the data in this Table, it can be seen that the highest producing clones from pPA003 are almost twice high as the clones derived from pPA005. As seen in Figure 7, the major difference between pPA003 and pPA005 is the inclusion from the normal t-PA gene of the intron that occurs within the pre- and pro- sequences of t-PA. This intron sequence could have enhancer-like activities as has been shown for the intron within immunoglobulin gene. Moreover, the

immunoglobulin enhancer sequence within the intron functions in a tissue-specific manner. Banerji, J., Olson, L., Schaffner, W. Cell 33: 729-740 (1983). It is possible that the sequences within the t-PA intron function in a similar manner.

TABLE III

Assayed Results of Selected Clones

Plasmid	#clones Assayed	Expression Levels						
		0-1.0	1.1-1.5	1.6-2.0	2.1-2.5	2.6-3.0	3.1-3.5	>3.5
pPA002	5	4	1	0	0	0	0	0
pPA003	9	1	1	1	3	0	1	2
pPA005	5	1	0	0	1	0	3	0
pPA007	2	0	1	1	0	0	0	0
pPA008	1	0	1	0	0	0	0	0
pPA014	7	2	5	0	1	0	0	0

* Expression levels are expressed as the ratio of the t-PA expresison of the clones vs. t-PA expression of parental CHL-1 cells. This value ranged from 46.7 to 88 ug/cell/day in three assays. The results of one or two assays of each clone have been averaged.

TABLE IV

COMPARISON OF EXPRESSION LEVELS OF CLONES GENERATED WITH pPA003 AND pPA005 TRANSFORMED INTO CHL-1

Clone Designation	Plasmid	Micro Units/ Cell/Day	Factor Enhanced
CHL-1	None	150 + 0.6	1.0
B11-3-C11	pPA003	780 + 3.7	4.6 + 0.5
B11-3-C8		730 + 2.8	4.7 + 0.2
51B3		650	4.1
B18-4-3	pPA005	400 + 0.8	2.5 + 0.4
53-3G3		380	2.4
B18-4-7		296	2.1
S4-A5		300	2

Enhancement factors were calculated by dividing the production level of the clone by the level of the parental cell line for each experiment. Production of t-PA from 10 cells plated in 1 ml medium for 72 h in a 24 well plate was measured by fibrin plates.

In one experiment, 10_6 CHL-1 cells were transfected with pPA003, 168 clones were generated, of which 73 were screened for t-PA expression levels as described above. The pPA003 clones had a range of t-PA expression levels ranging from 0.01 to 0.04 pg/cell/day. The percentage of clones producing increased levels of t-PA (32%) was lower than that observed for the hGH transfected clones of Example 1.

Several of the clones with the highest levels of t-PA expression were expanded and further analyzed for mRNA content. When the total cellular RNA was analyzed by a Northern blot analysis (Maniatis, T., Supra) two species of t-PA specific mRNA were detected, one corresponding to the host cell's native t-PA mRNA and the other to a dicistronic mRNA containing both the t-PA and neo gene sequences. The results of this Northern analysis is shown in Figure 8. This demonstration of a dicistronic mRNA species within the transfected clones indicates that one of the expression plasmids of this invention, pPA003, had successfully introduced an additional gene copy for t-PA into the CHL-1 clones. Moreover, since this plasmid contained sequences derived from genomic DNA which contains introns, the expression vector of this example system was able to process introns and transcribe active mRNA.

Several of the clones generated by t-PA expression plasmids containing other promotors were also analyzed for t-PA mRNA expression. The clones used in this analysis were the highest t-PA producing clone isolated from each promotor. The RNA samples were analyzed using Northern blot hybridization as described above. In this case the filter was hybridized with a t-PA and neo probes. In the case of hybridization with t-PA specific probe (Figure 9) two species of mRNA are detected only from clones generated with the LTR promotor. When this filter was then analyzed for neo RNA sequences (Figure 9) only the dicistronic mRNA

species contained neo sequences. The neo specific sequences in other clones were not detected by hybridization with this probe. (The amount of neo enzyme that is required for growth in G418 is very low and it is often difficult to detect neo mRNA in cells transfected with pneo 5, Figure 9, lane 8). The finding of stable production of dicistronic mRNA in cells transfected with plasmids containing the LTR promotor is consistent with the results of Table III in which the clones which produce the highest levels of t-PA following many cell divisions were generated with LTR containing plasmids.

In order to determine if one of the two mRNA species identified by the Northern analysis was initiated within the LTR promotor, an S1 nuclease analysis was carried out. To prepare the radioactive DNA probe, plasmid pneo5 was digested with Bgl II and EcoRI. The 2kb fragment was isolated and labeled with gamma- 32 p-ATP and polynucleotide kinase. This fragment was mixed with RNA obtained from the clone M5A6 and hybridized. The noncomplementary regions of the DNA probe were digested with S1 nuclease, and the protected DNA fragment was analyzed on a 5% acrylamide gel in 7M urea (Maniatis et al, supra). An autoradiogram of this gel analysis is shown in Figure 10. When mRNA from M5A6 is hybridized with the LTR probe, only one radioactive band is seen following digestion with S1 nuclease. The size of this protected fragment (275 bp) is consistent with initiation of mRNA transcription within the LTR promotor sequences. A similar band is not seen when the probe is hybridized with CHL-1 RNA.

The DNA of several clones derived from transfection of pPA003 into CHL-1 cells was analyzed by Southern blot. The DNA was digested with Cla I which has two recognition sites in pPA003 on either side of the LTR region. The DNA was electrophoresed through an agarose gel and transferred to a nylon filter membrane. This filter was hybridized with a radioactive probe

obtained from a Bgl II digest of the t-PA cDNA from pPA003. The hybridization pattern of the cloned DNA is seen in Figure 11. Only three of the four clones in this analysis were producing high levels of t-PA; the fourth clone (M5H6) produced high levels when it was originally isolated but the activity had dropped to background levels when the DNA was isolated.

In Figure 11, only one primary band is seen following digestion with Cla I. This Cla I fragment is the same size as the major Cla I fragment in the plasmid. This fragment is present in the three high t-PA producing clones isolated from this experiment. The filter containing the Cla I digested DNA was also hybridized with a probe containing sequences from the neomycin resistance gene. This probe also hybridizes to the 8 kb Cla I fragment that was complementary to the t-PA probe. The result of this analysis suggests that in the three highest producing clones isolated in this experiment, integration of the transfected plasmid occurred within the LTR region and thus did not interrupt the 8 kb Cla I fragment. Other high t-PA producing clones have been analyzed in the same manner and all of the DNA digestion patterns are consistent with integration within the LTR region.

Example 11

Examples 11 through 14 are summarized in Table V. The pPA003 transfection plasmid of Example 10 was used to transfect C_{127} cells, a mouse embryonic breast cell line, and clones resistant to G418 were selected. The G418 resistant clones numbered 44, of which 40 were screened for increased expression of t-PA, where the t-PA expression levels ranges from 0 to 0.1 pg/cell/day.

Example 12

The pPA003 transfection plasmid of Example 10 was used to transfect L-TK$^-$ cells, a mouse fibroblast cell line, and clones resistant to G418 were selected. The G418 resistant clones numbered 42, of which 36 were

screened for increased expression of t-PA, where the t-PA expression levels ranges from 0 to 0.1 pg/cell/day.

Example 13

Transfection plasmid pPA201 was constructed and is shown in Figure 12. The t-PA sequences described in Example 10 were introduced into the eukaryotic expression vector containing the mouse dihydrofolate reductase (dhfr) gene for a marker (pSV-dhfr, Example 7). Again, the t-PA gene was inserted between the promotor and the selection marker. This plasmid was introduced into Chinese Hamster Ovary (CHO) cells, a cell line which lacks the normal dhfr gene. Clones were selected for growth in F12 medium lacking hypoxanthine, glycine, and thymine. The clones were assayed for t-PA activity by ELISA. The assays showed that several of the clones were secreting t-PA at level of 2 to 8 x 10^{-3} pg/cell/day. These clones were then placed in amplification medium containing 0.2 uM MTX and subclones were isolated. When these subclones were assayed for t-PA production by ELISA, all of the subclones were shown to secrete substantially more t-PA (up to 165 times more) than their parental clones. This increased t-PA production was probably the result of amplification of the dhfr gene in response to the increased levels of MTX as in Example 9. Thus, it is possible to amplify two genes that are linked by a common promotor.

Example 14

The pPA401 expression plasmid is shown in Figure 13. To construct this plasmid, the neo sequences were deleted from and replaced with sequences encoding the herpes simplex virus thymidine kinase gene. Plasmid pPA401 was used to transfect L-TK^- cells, a mouse fibroblast cell line that lacks thymidine kinase. Clones were selected by their ability to grow in the presence of methotrexate, hypoxanthine and thymidine. Sixty one (61) clones were obtained by this procedure, and all were screened for t-PA activity. Only four of these

clones expressed human of up to 0.01 pg/cell/day. Moreover, the number of clones derived from pPA401 expressing t-PA (4/61) is much lower than that seen with pPA003 into L-TK$^-$ cells (29/40), and the expression levels are ten-fold lower (see Table V).

Recombinant DNA described in Example 10, in an E.coli host was deposited with the American Type Culture Collection (ATCC), Rockville, Md. as follows: MH1/pPA003 was deposited on January 13, 1987 and assigned ATCC accession number 67293.

TABLE V

Example	Plasmid	Host Cell	Agent	Transfection Efficiency (x10-4)	tPA+ clones/ Total Analyzed	Expression Levels pg/cell/day
10	pPA003	CHL-1	G418	0.07	73/168	0.01-0.4
11	pPA003	C127	G418	1.1	29/40	0 - 0.1
12	pPA003	L-TK$^-$	G418	2.1	7/36	0 - 0.1
13	pPA201	CHO_dhfr$^-$	MTX	.01	2/3	0 - .011
14	pPA401	L-TK$^-$	MTX thymidine and hypoxanthine	3.0	4/61	0 - 0.01

Erythropoetin

Erythropoetin ("EPO") is a glycoprotein that stimulates the formation of red blood cells and has been shown to be effective in early clinical trials in the treatment of dialysis patients.

For each transfection example for EPO described hereinbelow, 5 ug of transfection plasmid DNA was added to a monolayer of approximately 2 x 10^5 host cells in 6-well plates. In each of the following examples, the selective phenotype confers resistance to G418 at various concentrations. This concentration was individually determined for each cell line. Clones which demonstrated G418 resistance were identified 3 to 4 weeks after transfection. The results of these examples are summarized in Table VI below.

Example 15

Transfection plasmid pEP1 is shown in Figure 14. It was constructed by insertion of a DNA fragment encoding EPO, that was obtained from a human genomic library in lambda phage, into a derivative of the pneo5 disclosed in Example 1. In this example, the pneo5 plasmid was first modified by the insertion of a polylinker which carries sites for Pvu I, Sac II, Mlu I, Apa I, Hind III and Bcl I. A 4 kb Apa I fragment containing the EPO coding sequences was inserted into the Apa I site in the polylinker region. The resultant plasmid pEP1 carries the erythropoetin gene between the promotor and the resistance marker. The expected mRNA will be dicistronic encoding EPO in the 5' cistron and neo in the downstream 3' cistron.

An EPO gene was also constructed synthetically and lacked the introns of the genomic fragment. Synthetic DNA oligomers were annealed together in 4 groups of 8 and 1 group of 6 oligomers. Each of these had unique restriction sites at their ends and were cloned into a pUC plasmid cut with the cognate enzyme. Plasmid pF which contains the complete EPO gene flanked by Bam

HI sites, was cut with Bam HI and the EPO gene was ligated into the Bgl II site of pneo5 creating pEP2. Plasmid pEP2 encodes EPO and neo on a dicistronic mRNA with the EPO gene lacking introns and therefore equivalent to a cDNA gene.

Example 16

The following examples deal with the transfection of different host cells. This data is summarized in Table VI.

The above-described plasmid of Example 15 was used to transfect CHL-1 cells and clones resistant to G418 were isolated. Erythropoetin production was detected by the Krystal assay which stimulates ^{3}H thymidine uptake into spleen cells. The number of isolated clones numbered 710, all of which produced EPO in the range of 0.04 to 0.5 pg/cell/day. The identity of the EPO was confirmed by its in vivo (exhpoxic mouse assay) activity and by its recognition in a radio immune assay. However, the number of units determined by the in vivo assay using authentic human EPO was only 8% of the number of units determined by the in vitro Krystal assay using the same standard.

Example 17

The pEP2 transfection plasmid of Example 15 was used to transfect 2 X 10^{6} CHL-1 cells and clones resistant to G418 were isolated. The G418 resistant clones numbered 134, of which 24 were screened for expression of EPO. The EPO production by these clones was less than 0.2 units/ml and, therefore, below detection by the Krystal assay.

Example 18

The pEP1 transfection plasmid of Example 15 was used to transfect 2 x 10^{5} cells of a human embryonic kidney cell line 293 and clones resistant to G418 were isolated. The G418 resistant clones numbered 96, of which 24 were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from

0 to 0.28 pg/cell/day. The identity of EPO was confirmed by its in vivo activity. The number of units obtained by the in vivo assay was 2.3% of the number obtained from the in vitro assay using the same human standard.

Example 19

The pEP1 transfection plasmid of Example 15 was used to transfect 2 x 10^5 MDBK cells, a bovine kidney cell line, and clones resistant to G418 were isolated. The G418 resistant clones numbered 25, of which 8 were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from 0 to 0.4 pg/cell/day. The identity of EPO was confirmed by its in vivo activity. The number of units obtained by the in vivo assay was 2.1% of the number obtained from the in vitro assay using the same human standard.

Example 20

The pEP1 transfection plasmid of Example 15 was used to transfect 2 x 10^5 C_{127} cells, a mouse embryonic breast cell line and clones resistant to G418 were isolated. The G418 resistant clones numbered 96, of which 24 were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from 0 to 0.54 pg/cell/day. The identity of EPO was confirmed by its in vivo activity. The number of units obtained by the in vivo assay was 17.4% of the number obtained from the in vitro assay using the same human standard.

Example 21

The pEP1 transfection plasmid of Example 15 was used to transfect 2 x 10^5 L-TK^- cells, a mouse fibroblast cell line, and clones resistant to G418 were isolated. The G418 resistant clones numbered 408, of which 24 were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from 0.08 to 0.34 pg/cell/day. The identity of EPO was confirmed by its in vivo activity. The number of units

obtained by the in vivo assay was 33.6% of the number obtained from the in vitro assay using the same human standard.

Example 22

The pEP2 transfection plasmid of Example 15 was used to transfect 2 X 10^5 L-TK$^-$ cells and clones resistant to G418 were isolated. The G418 resistant clones numbered 34, of which 24 were screened for expression of EPO. The EPO production by these clones was less than 0.2 units/ml and, therefore, below detection by the Krystal assay.

Example 23

The pEP1 transfection plasmid of Example 15 was used to transfect 2 x 10^5 COS-7 cells, a green monkey kidney cell line, and clones resistant to G418 were isolated. The G418 resistant clones numbered 84, of which 24 were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from 0 to 0.63 pg/cell/day. The identity of EPO was confirmed by its in vivo activity. The number of units obtained by the in vivo assay was 2.3% of the number obtained from the in vitro assay using the same human standard.

Example 24

The pEP1 transfection plasmid of Example 15 was used to transfect 1.5 x 10^6 CHO dhfr$^-$ cells, a Chinese Hamster Ovary cell line (CHO), and clones resistant to G418 were isolated. The 24 G418 resistant clones were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from 0 to 0.88 pg/cell/day. The identity of EPO was confirmed by its in vivo activity. The number of units obtained by the in vivo assay was 95% of the number obtained from the in vitro assay using the same human standard. This is the only cell line where we obtained an equivalence between in vitro and in vivo activity.

Example 25

The pEP2 transfection plasmid was used to transfect 3 X 10^5 CHO dhfr$^-$ cells, and clones resistant to G418 were isolated. The G418 resistant clones numbered 43, of which 36 were screened for expression of EPO. The EPO producing clones had EPO expression levels ranging from 0 to 0.25 pg/cell/day. This is the only cell line examined in which the pEP2 plasmid was successful in generating clones that produce EPO. These examples show that the number of transfectants isolated and their EPO production was improved when pEP1 was used rather than pEP2. Therefore, the EPO containing introns, is preferable to the EPO lacking introns.

Recombinant DNA and host cells described in Example 15 in an E.coli host were deposited with the American Type Culture Collection (ATCC), Rockville, Md. as follows:

MHI/pEPI was deposited on January 13, 1987 and assigned ATCC accession number 67292.

TABLE VI

Example	Plasmid	Host Cell	Transfection Frequency	Expression Levels pg/cell/day 10^{-2}	% Activity in vivo/ in vitro
17	pEP1	CHL-1	10^{-3}	4-50	8.1%
18	pEP2	CHL-1	2.2×10^{-4}	0	-
19	pEP1	293	5×10^{-4}	0-28	2.3
20	pEP1	MDBK	10^{-4}	0-40	2.1
21	pEP1	C127	5×10^{-4}	0-54	17.4
22	pEP1	LTK^-	2×10^{-3}	8-34	33.6
23	pEP2	LTK^-	8.5×10^{-4}	0	-
24	pEP1	COS^{-7}	4×10^{-4}	0-63	2.3
25	pEP1	CHO $dhfr^-$	2×10^{-5}	0-88	95.0
26	pEP2	CHO $dhfr^-$	3×10^{-5}	0-25	-

Although the foregoing invention has been described in some detail by way of illustration, for the purpose of clarity and understanding, it will be appreciated that numerous modifications may be practiced within the spirit and scope of the appended claims.

CLAIMS:

1. A method for producing a desired protein in a eukaryotic cell comprising:

(a) forming a eukaryotic dicistronic expression vector, which vector comprises:

(i) an intron-containing structural DNA sequence which codes for the desired protein, and

(ii) a second DNA sequence which codes for a selectable protein;

wherein the sequences are operably linked to one another and depend on one common promotor; and

(b) transfecting a eukaryotic cell with the vector and allowing this cell to grow under conditions favorable to the production of the selectable protein.

2. A method according to claim 1, wherein the desired protein is t-PA.

3. A method according to claim 2, wherein a portion of the structural DNA sequence is derived from genomic DNA and contains at least one intron of between about 1-2kb.

4. A method according to claim 2, wherein the intron is Intron A.

5. A method according to claim 1, wherein the desired protein is erythropoetin (EPO).

6. A method according to claim 1, wherein the structural DNA sequence coding for the desired protein is disposed between the promotor and the second DNA sequence.

7. A method according to Claim 1, wherein the promotor comprises a plurality of long term repeat regions of Harvey murine sarcoma virus.

8. A host cell which has been transfected with a eukaryotic dicistronic expression vector, which vector comprises:

(i) An intron-containing structural DNA sequence which codes for a desired protein, and

(ii) a second DNA sequence which codes for a selectable protein; and

wherein the sequences are operably linked to one another and depend on one common promotor.

9. Recombinant DNA material comprising:

(a) an intron-containing structural DNA sequence which codes for a desired protein, and

(b) a second DNA sequence which codes for a selectable protein, wherein the sequences are operably linked to one common promotor.

10. A method for producing t-PA in a eukaryotic cell comprising:

(a) forming a eukaryotic expression vector that comprises a cDNA sequence and a genomic DNA sequence with at least one intron between about 1-2kb wherein said sequences encode t-PA; and

(b) transfecting the cell with the vector.

11. A method according to Claim 10, wherein the intron is Intron A.

12. A hybrid t-PA gene comprising a cDNA sequence and a genomic DNA sequence with at least one intron between about 1-2kb wherein said sequences encode t-PA.

13. The hybrid t-PA gene of Claim 12, wherein the intron is Intron A.

14. An expression vector capable of stable maintenance in a cell culture comprising a hybrid t-PA gene according to Claim 12.

15. An expression vector capable of stable maintenance in a cell culture comprising a hybrid t-PA gene according to Claim 13.

16. A cell in culture containing an expression vector according to Claim 14.

17. A cell in culture containing an expression vector according to Claim 15.

18. A cell line and recombinant DNA material contained therein which is identified as MH1/pPA003 (ATCC 67293).

19. A cell line and recombinant DNA material contained therein which is identified as MH1/pPEPI (ATCC 67292).

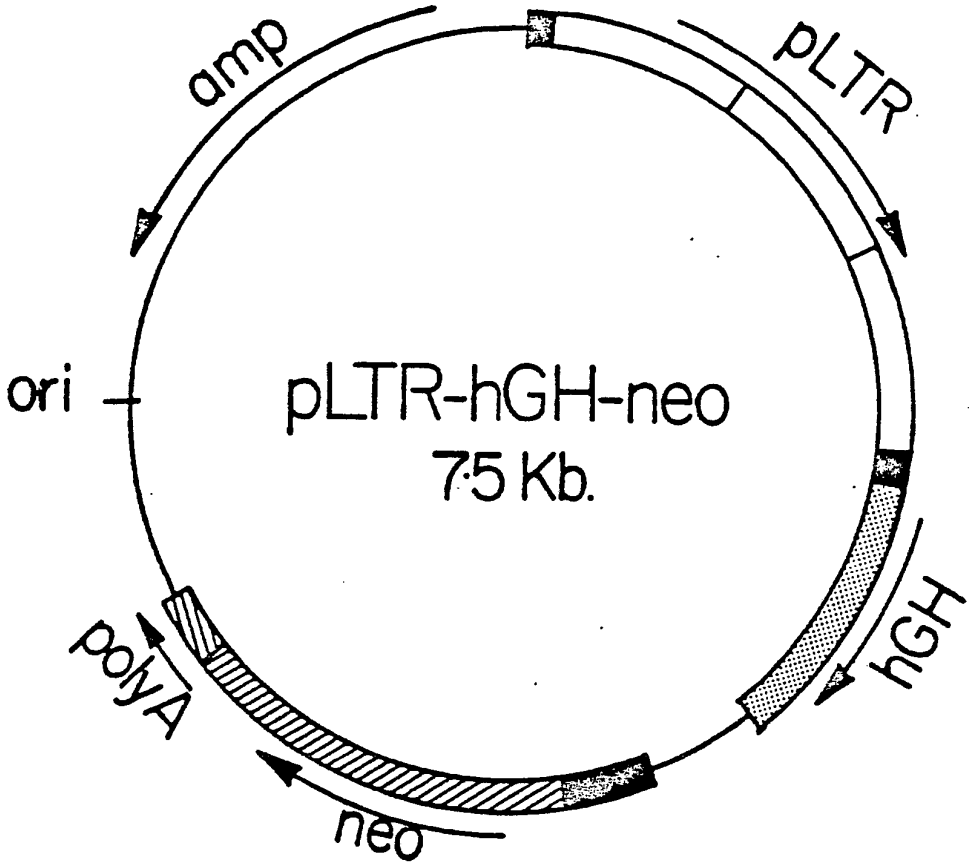


FIGURE 1

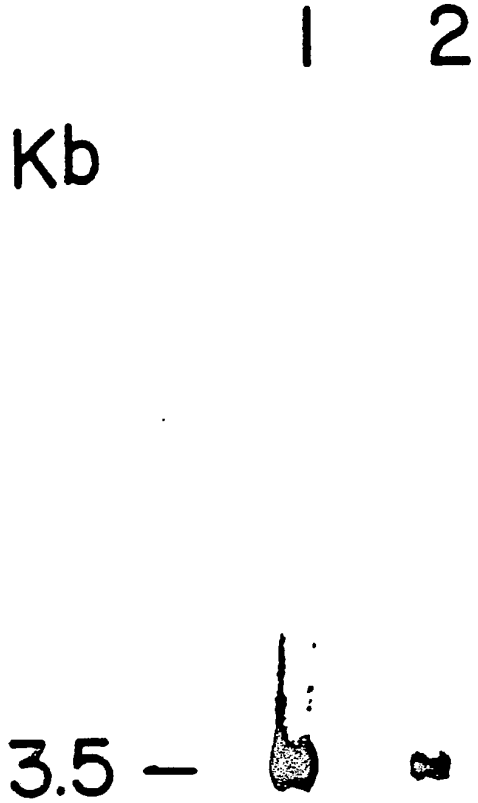
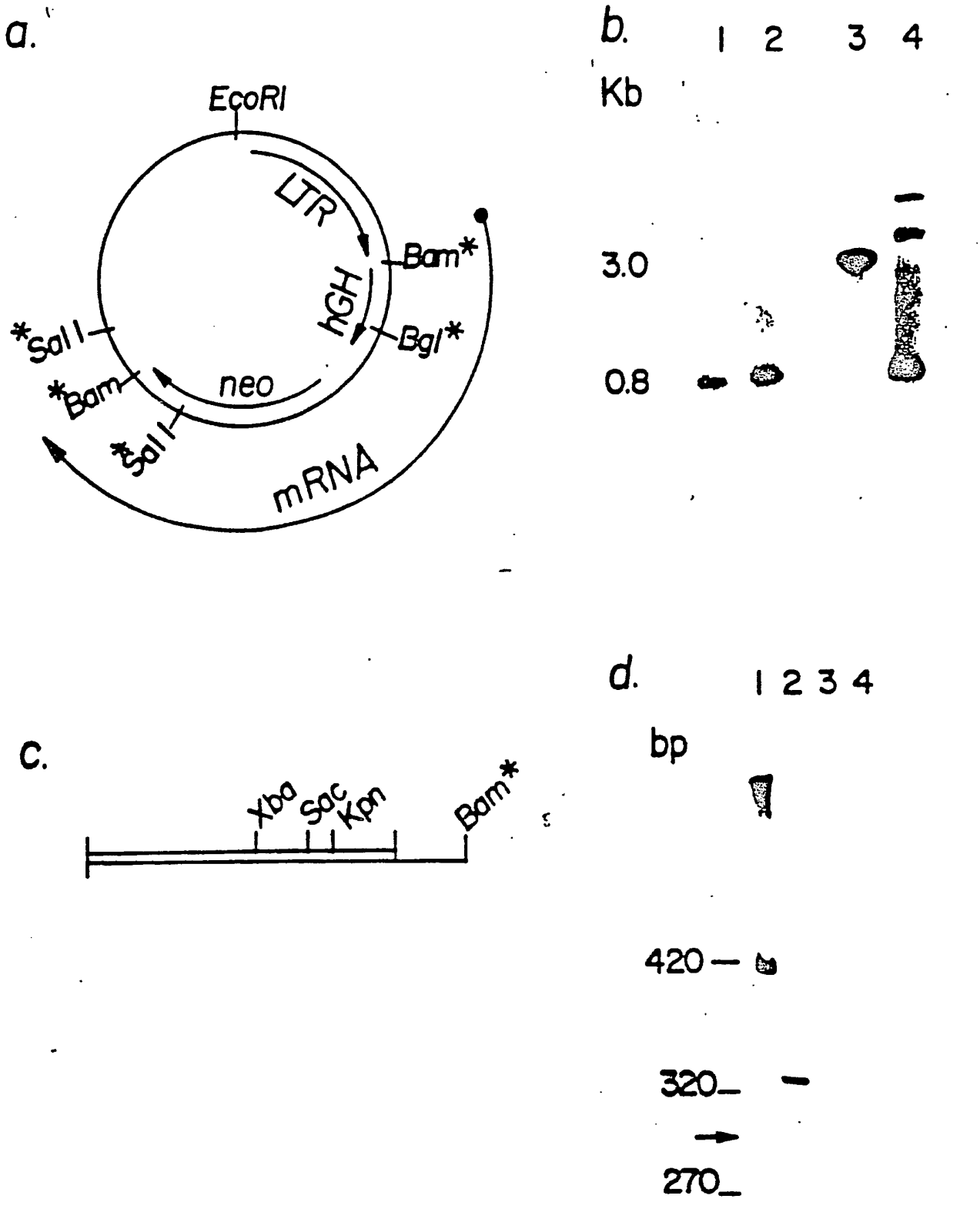


FIGURE 2

FIGURE 3


a.
EcoRI
LTR
Bam*
hGH
Bgl*
*Sal I
*Bam
neo
*Sal I
mRNA
b.
1 2 3 4
Kb
3.0
0.8
c.
Xba
Sac
Kpn
Bam*
d.
1 2 3 4
bp
420
320
270

SOUTHERN BLOT ANALYSIS OF CLONE D4

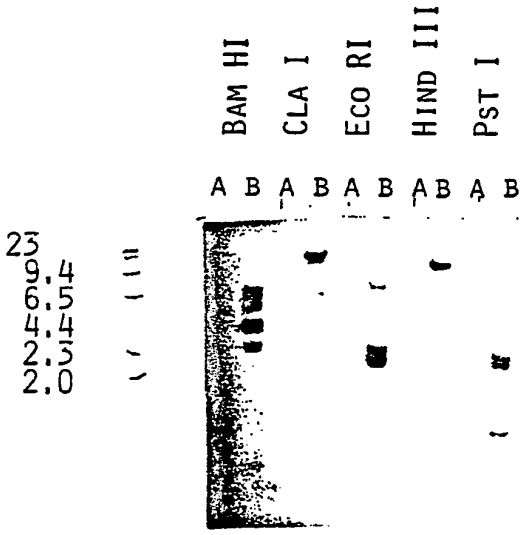


FIGURE 4

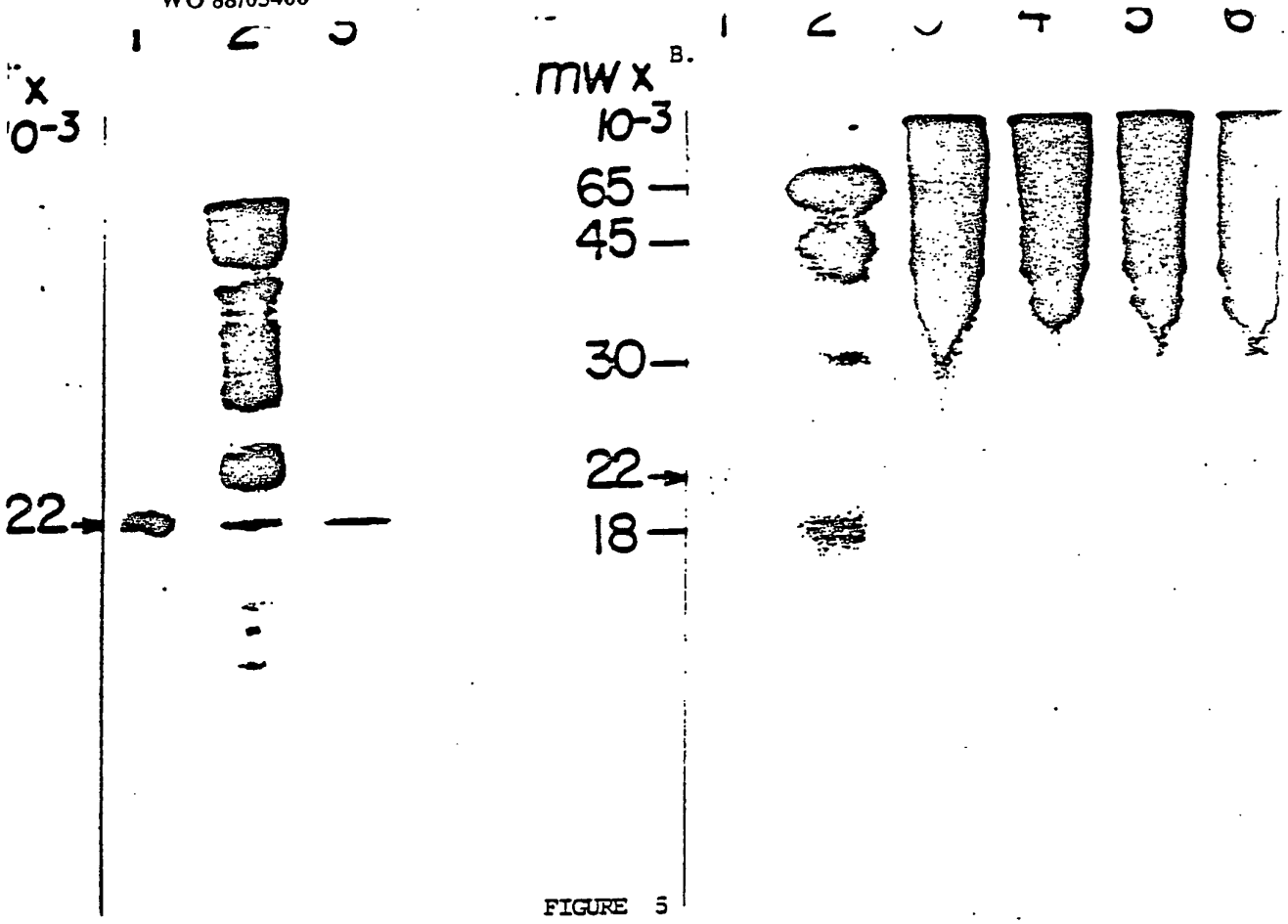
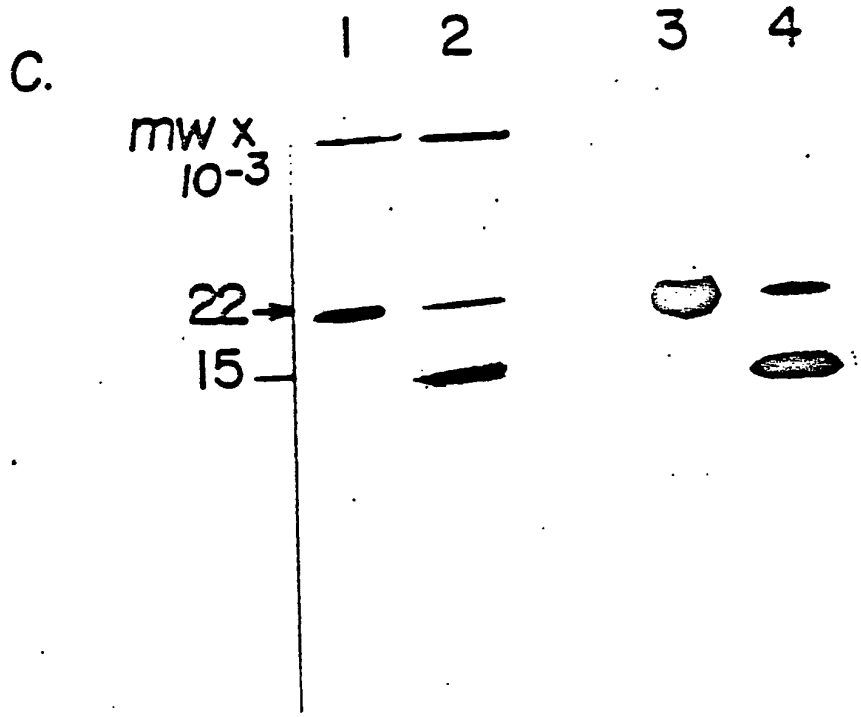


FIGURE 5



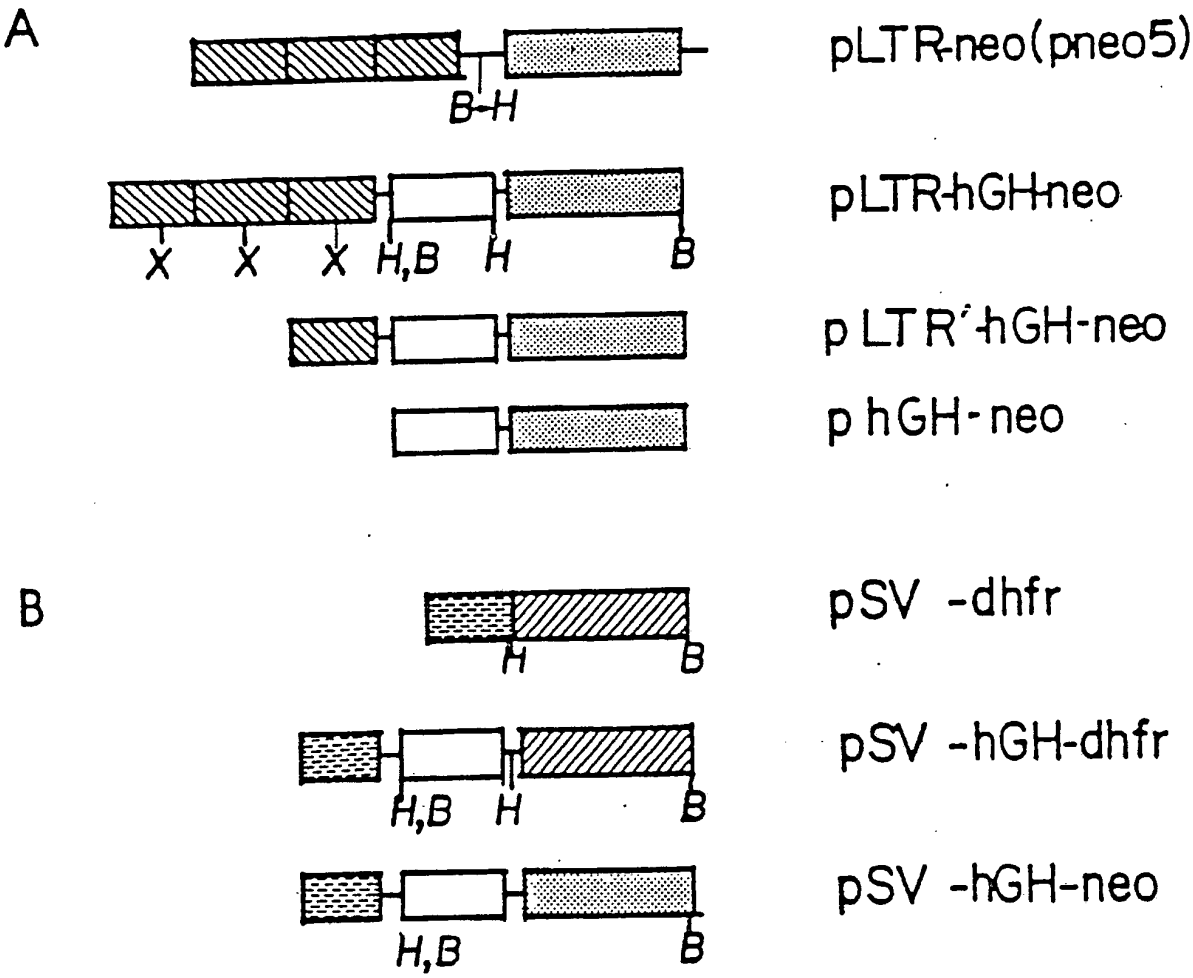

A
B-H
pLTR-neo(pneo5)
X X X H,B H B
pLTR-hGH-neo
pLTR-hGH-neo
phGH-neo
B
H B
pSV -dhfr
H,B H B
pSV -hGH-dhfr
H,B B
pSV -hGH-neo

FIGURE 6

Figure 7b

LIGATION OF SYNTHETIC OLIGOMERS
TO FORM AMINO TERMINAL (PRE-PRO) REGION OF t-PA GENE

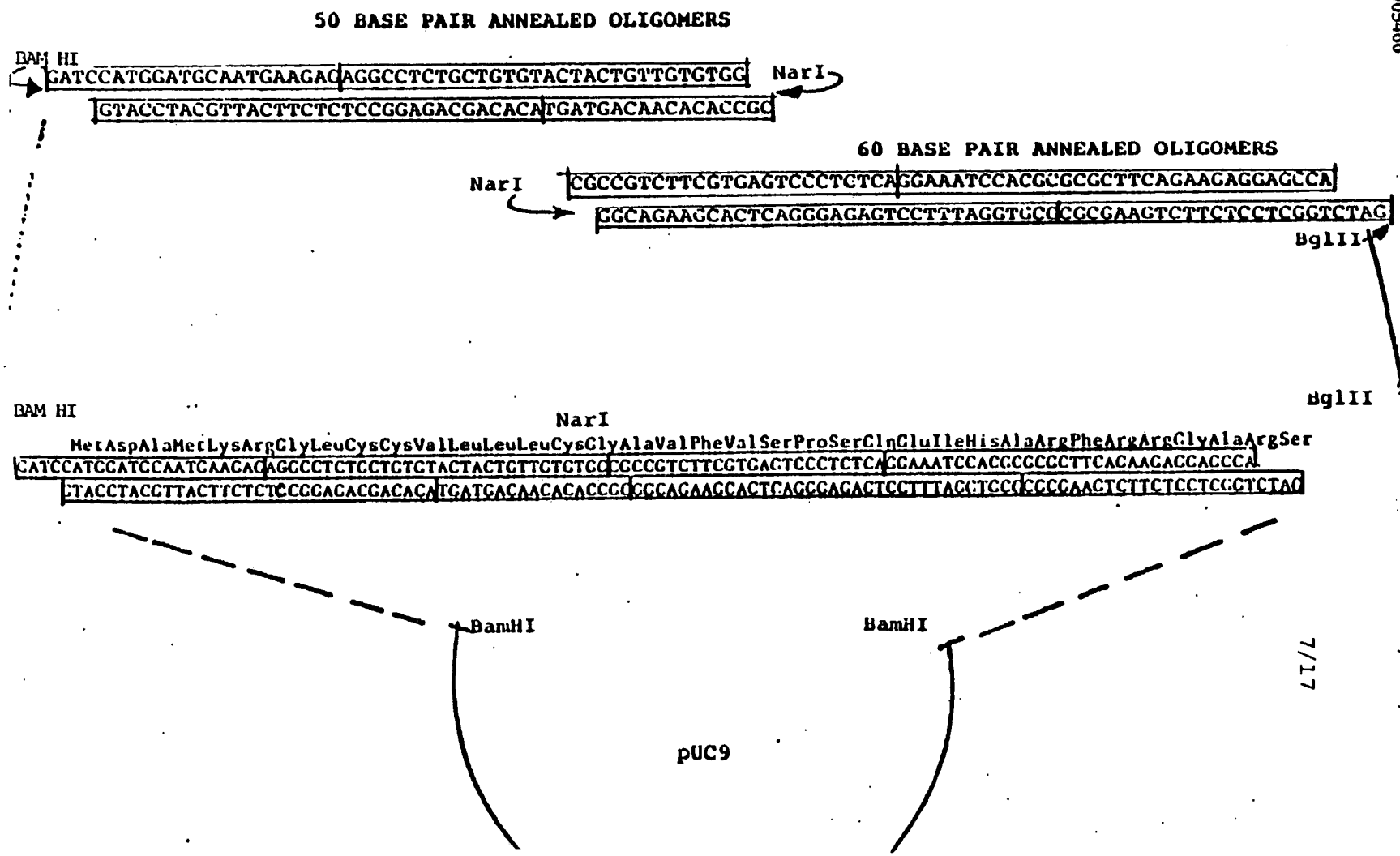


Figure 7c

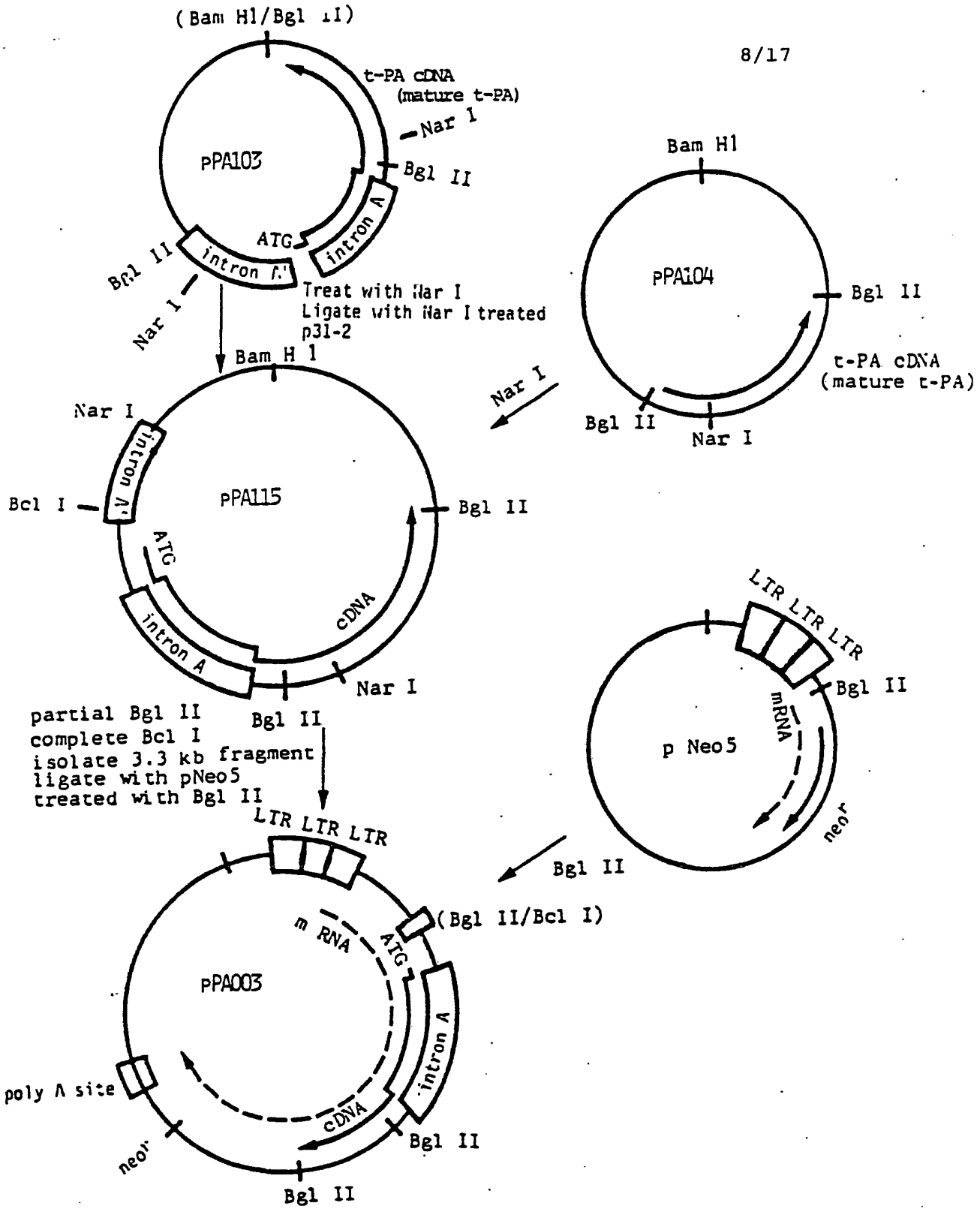
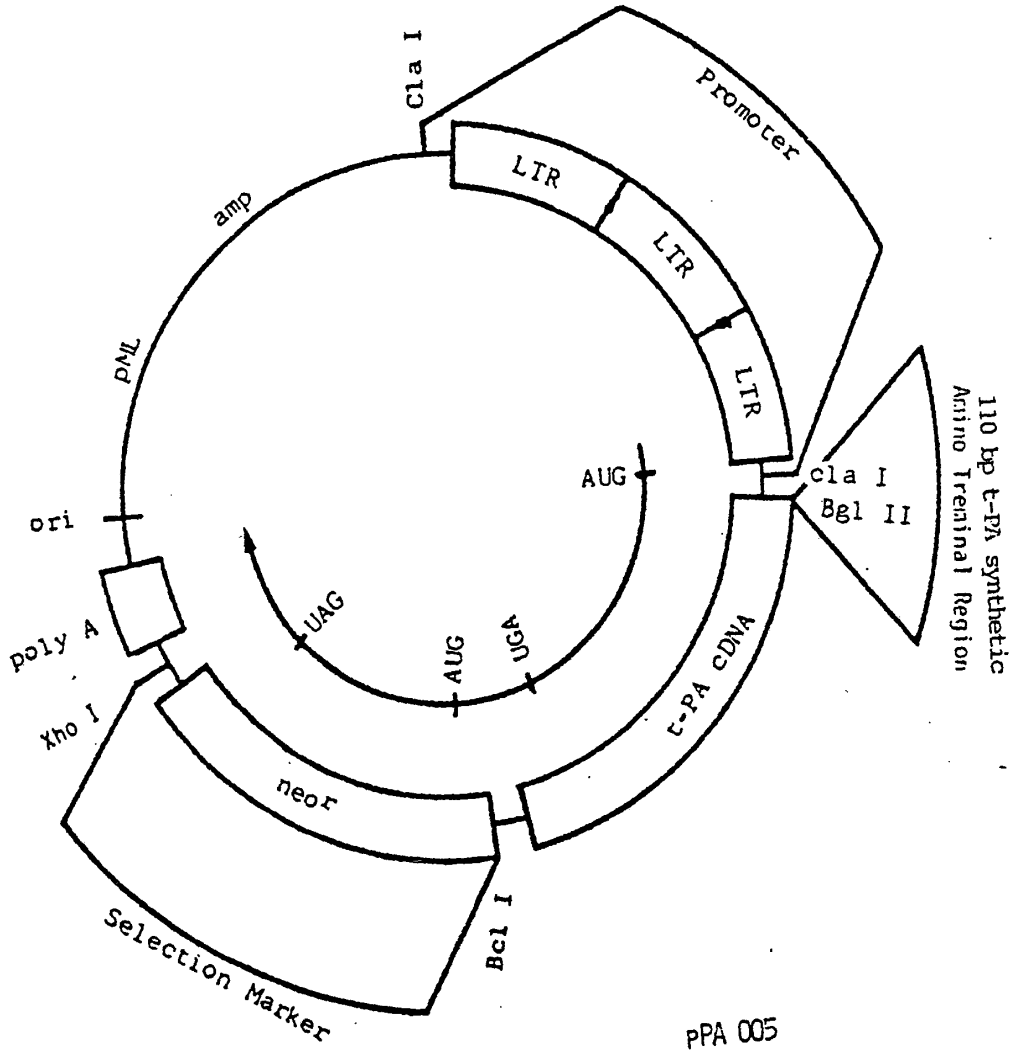

(Bam H1/Bgl II)
t-PA cDNA
(mature t-PA)
Nar 1
pPA103
Bgl II
ATG
intron A
Bgl II
intron A
Nar 1
Treat with Nar I
Ligate with Nar I treated
p31-2
Bam H1
pPA104
Bgl II
t-PA cDNA
(mature t-PA)
Bgl II
Nar I
Nar I
Bam H 1
Nar I
intron A
Bcl I
pPA115
Bgl II
ATG
cDNA
intron A
Nar I
Bgl II
partial Bgl II
complete Bcl I
isolate 3.3 kb fragment
ligate with pNeo5
treated with Bgl II
LTR LTR LTR
Bgl II
p Neo5
mRNA
neor
Bgl II
LTR LTR LTR
(Bgl II/Bcl I)
m RNA
ATG
pPA003
intron A
poly A site
cDNA
neor
Bgl II
Bgl II

Figure 7d


Cla I
Promoter
LTR
LTR
LTR
amp
pML
ori
poly A
Xho I
Selection Marker
neor
Bcl I
t-PA cDNA
cla I
Bgl II
110 bp t-PA synthetic
Amino Treminal Region
AUG
UAG
AUG
UGA
pPA 005

NORTHERN ANALYSIS OF RNA FROM CLONES GENERATED WITH pPA003 AND HYBRIDIZED WITH A T-PA PROBE

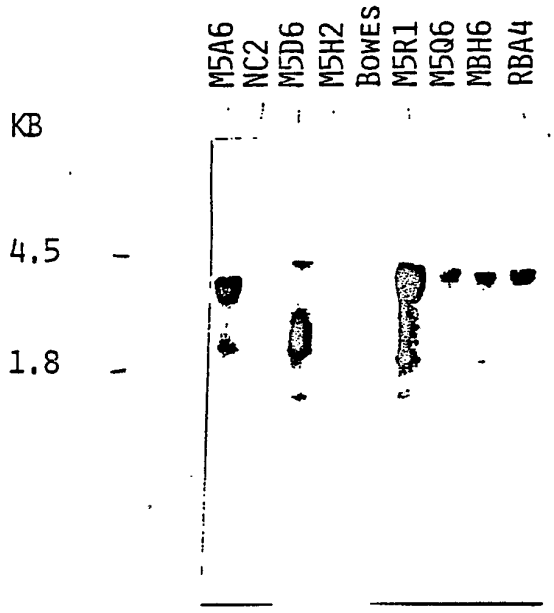


FIGURE 8

NORTHERN ANALYSIS OF CLONES GENERATED WITH T-PA EXPRESSION PLASMIDS THAT CONTAIN DIFFERENT PROMOTERS

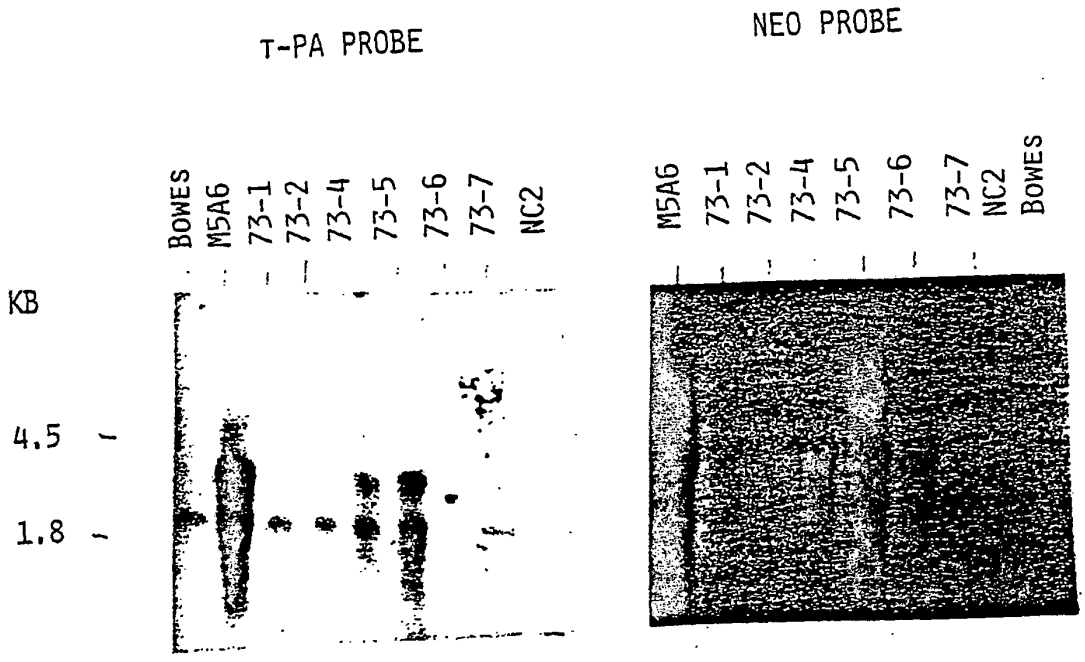
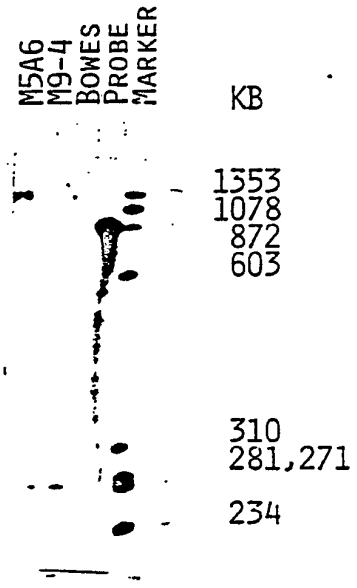


FIGURE 9

S1 ANALYSIS OF RNA FROM M5A6



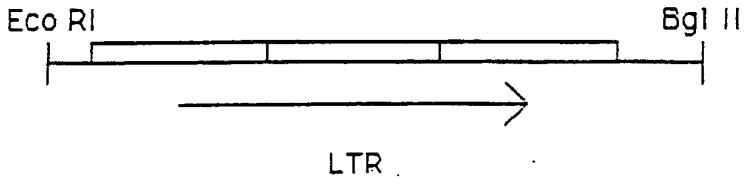


FIGURE 10

DIGESTION OF DNA WITH CLA I DERIVED FROM CLONES GENERATED WITH pPA003

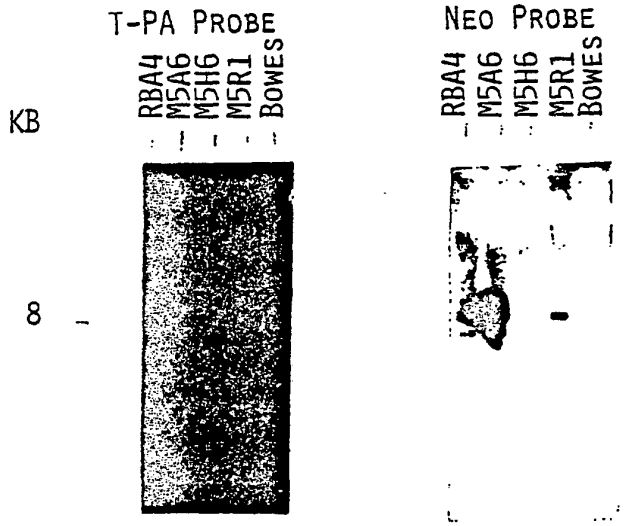


FIGURE 11

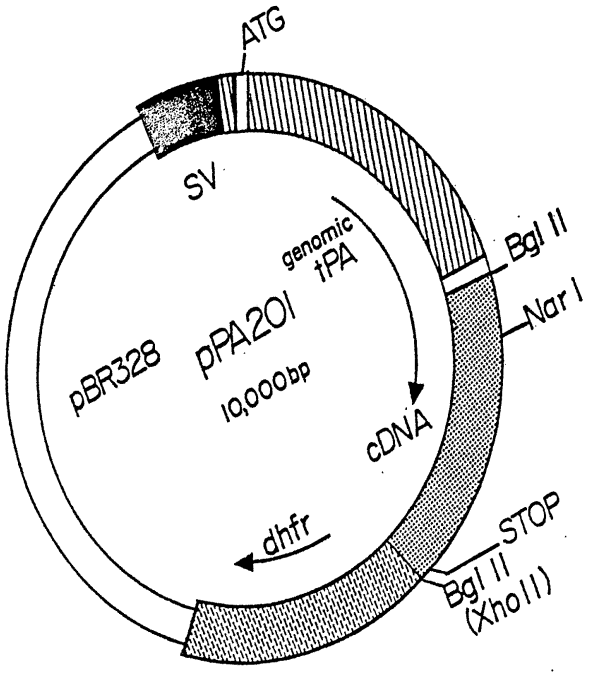

ATG
SV
genomic
tPA
Bgl II
Nar I
pBR328
pPA201
10,000bp
cDNA
dhfr
STOP
Bgl II
(Xho II)

FIGURE 12

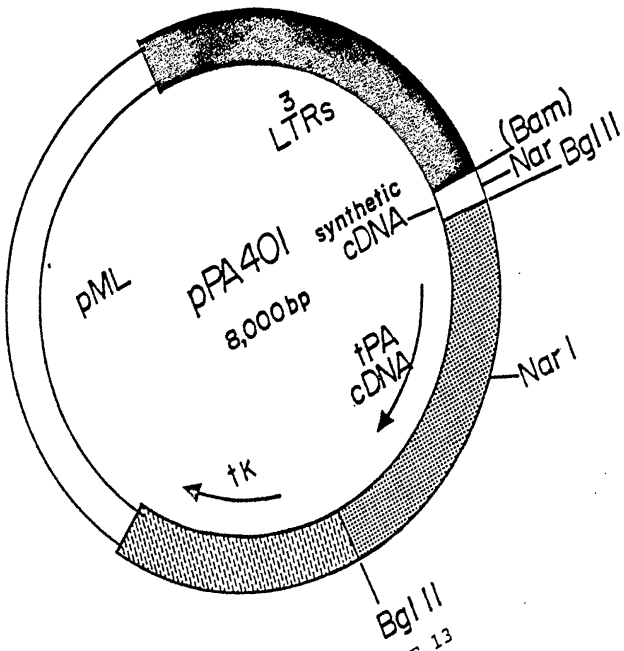

3
LTRs
(Bam)
Nar
Bgl II
synthetic
cDNA
pML
pPA401
8,000bp
tPA
cDNA
Nar I
tK
Bgl II

FIGURE 13

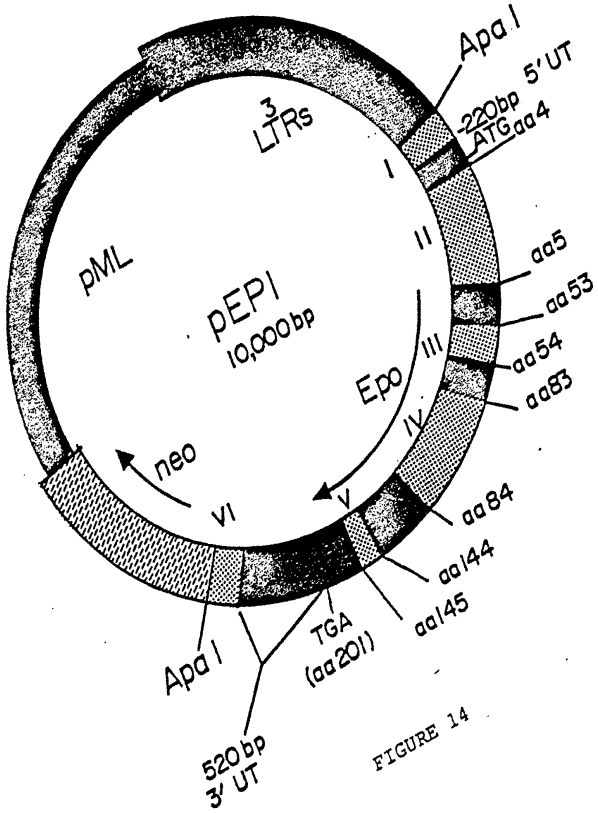

Apa I
220bp 5' UT
ATG aa 4
3 LTRs
I
II
pML
pEP1
10,000 bp
aa 5
aa 53
aa 54
aa 83
III
Epo
IV
neo
VI
V
aa 84
aa 144
aa 145
TGA
(aa 201)
Apa I
520 bp
3' UT

FIGURE 14

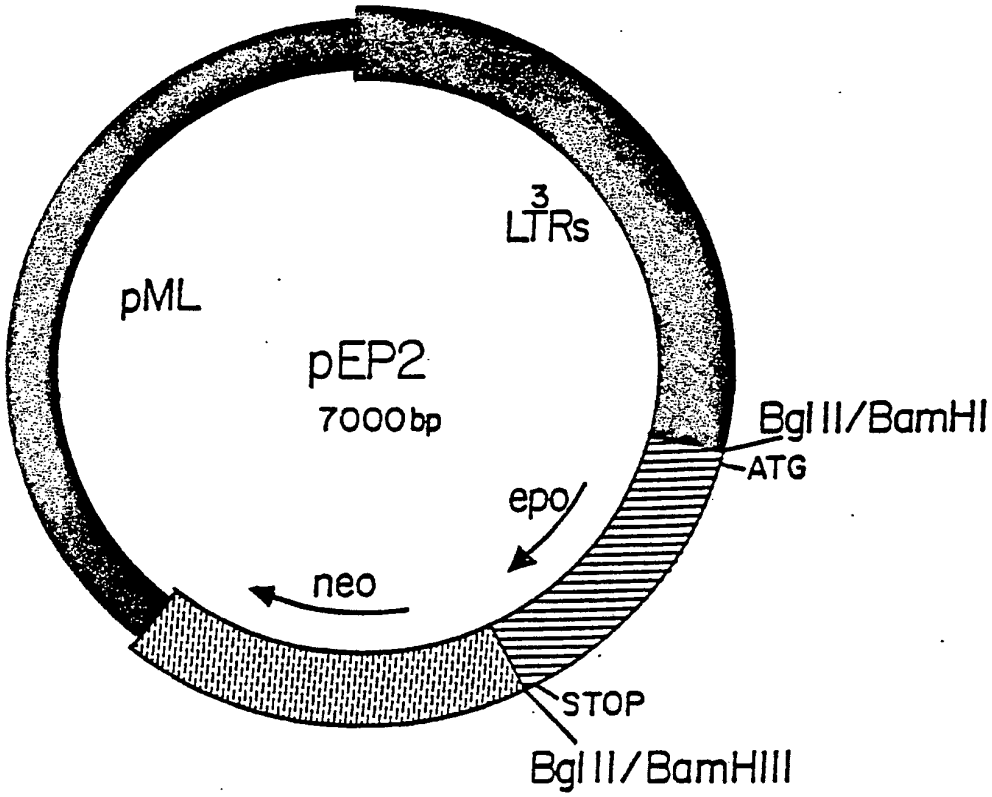

3
LTRs
pML
pEP2
7000bp
BglII/BamHI
ATG
epo
neo
STOP
BglII/BamHIII

FIGURE 15

REVISED VERSION *

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C12N 15/00, C12P 21/02
C12N 9/64

A3

(11) International Publication Number: WO 88/ 05466

(43) International Publication Date: 28 July 1988 (28.07.88)

(21) International Application Number: PCT/US88/00109

(22) International Filing Date: 15 January 1988 (15.01.88)

(31) Priority Application Number: 003,611

(32) Priority Date: 15 January 1987 (15.01.87)

(33) Priority Country: US

(71) Applicant: CODON [US/US]; 213 East Grand Avenue, South San Francisco, CA 94070 (US).

(72) Inventors: MORSER, Michael, John ; 3964 - 20th Street, San Francisco, CA 94114 (US). CASHION, Linda ; 219 Kelton Avenue, San Carlos, CA 94070 (US).

(74) Agent: WEBER, Ellen, L.; Townsend and Townsend, One Market Plaza, 2000 Steuart Tower, San Francisco, CA 94105 (US).

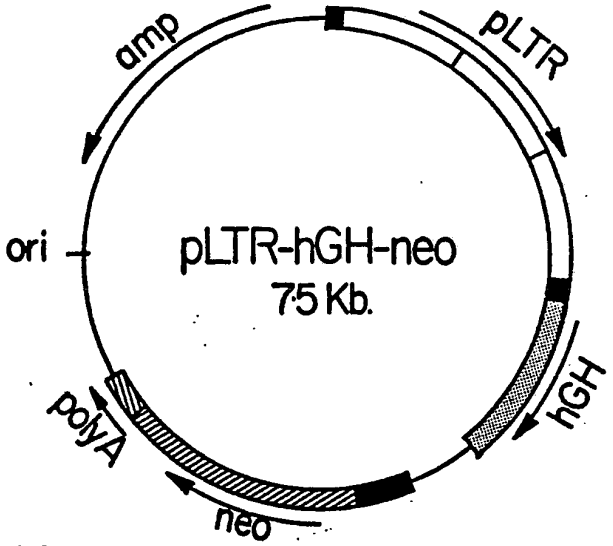
(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With a revised version of international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
11 August 1988 (11.08.88)

Date of publication of the revised version of the international search report:
3 November 1988 (03.11.88)

(54) Title: TANDEM GENE EUKARYOTIC EXPRESSION VECTORS



(57) Abstract

Compositions and methods for expressing gene products of interest in eukaryotic cells by transforming eukaryotic cells with recombinant DNA expression vectors which contain a DNA sequence coding for a gene product of interest together with a DNA sequence coding for a selectable phenotype. Both the gene of interest and the selectable gene are transcribed from a single eukaryotic promotor. Preferably, the recombinant DNA constructed is such that the DNA sequence coding for the gene product of interest is disposed between the eukaryotic promotor and the DNA sequence coding for the selectable phenotype. In an expression vector thus constructed, the gene product of interest is necessarily expressed by those cells which express the selectable phenotype. One clone of such transformed eukaryotic cells contains dicistronic mRNA which encodes for both human tissue plasminogen activator and the selectable phenotype. Another clone contains dicistronic mRNA which encodes for both erythropoetin and the selectable phenotype. Also disclosed are recombinant genes including an intron. The recombinant genes may be hybrids including both genomic DNA and cDNA.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

REVISED VERSION

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 88/00109

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00; C 12 P 21/02; C 12 N 9/64

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X	EP, A, 0227064 (BOEHRINGER MANNHEIM GmbH) 1 July 1987 see page 2, line 49 - page 3, line 14 --	10,12,14, 16
A	Proceedings of the National Academy of Sciences of the USA, volume 81, no. 17, September 1984, (Washington, D.C., US), T. Ny et al.: "The structure of the human tissue-type plasminogen activator gene: correlation of intron and exon structures to functional and structural domains", pages 5355-5359 --	
X	WO, A, 86/05514 (CHIRON CORP.) 25 September 1986 see the whole document	10-17
Y	--	2-4
X	EP, A, 0117058 (GENENTECH, INC.) 29 August 1984 ./.	1,6.8,9

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
23rd September 1988	12 OCT 1988
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	P.C.G. VAN DER PUTTEN

Form PCT/ISA/210 (second sheet) (January 1985)

FURTHER INFORMATION CONTINUED FROM THE SE OND SHEET

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers ..... because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim numbers ....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

1. Claims 1-9, 19

2. Claims 10-18

Please refer to Form PCT/ISA/206 sent to you on May 9th, 1988.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☒ No protest accompanied the payment of additional search fees.

III. DOCUMENTS CONSIDERED T BE RELEVANT (CONTINUED FR M THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
Y	see claims 1,13,18	2-5,7
X	EP, A, 0191606 (ELI LILLY AND CO.) 20 August 1986 see page 17, lines 7-11; figure 4	1,6
Y	P.H. Pouwels et al.: "Cloning vectors: A laboratory manual", 1985, Elsevier, (Amsterdam, NL), see page VIII-A-c-i-4	7
Y	WO, A, 85/02610 (KIRIN-AMGEN, INC.) 20 June 1985 see figure 3	5
A	Journal of Molecular and Applied Genetics, volume 1, no. 4, August 1982, Raven Press, (New York, US), P.J. Southern et al.: "Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter", see pages 327-341 cited in the application	

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. US 8800109 SA 20608

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 03/10/88
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 227064	01-07-87	DE-A- 3545126	25-06-87
		WO-A- 8703904	02-07-87
		AU-A- 6833387	15-07-87
		JP-T- 62502942	26-11-87
		EP-A- 0250513	07-01-88
WO-A- 8605514	25-09-86	EP-A- 0215923	01-04-87
		JP-T- 62502585	08-10-87
EP-A- 0117058	29-08-84	AU-A- 2353184	26-07-84
		JP-A- 59173096	29-09-84
		US-A- 4713339	15-12-87
EP-A- 0191606	20-08-86	AU-A- 5325686	14-08-86
		JP-A- 61205487	11-09-86
WO-A- 8502610	20-06-85	EP-A- 0148605	17-07-85
		AU-A- 3746785	26-06-85
		JP-T- 61501627	07-08-86
		US-A- 4703008	27-10-87